The Extended Decision Model

by

Darren R. Christensen

A thesis submitted in fulfillment of the requirements for the degree of the

Doctor of Philosophy

in

Psychology

University of Canterbury

New Zealand

April 20 2009

Acknowledgements

I am extremely grateful to my supervisor Associate Professor Randolph Grace for the opportunity to be his student. His suggestions regarding my thesis were always very helpful and much appreciated. I would also like to thank the Head of Department Associate Professor Neville Blampied for his efforts on my behalf at pivotal moments during my postgraduate years. Finally, I would like to thank my wife Josephine, for her patience, understanding, and constant encouragement.

Contents Page

List of Equations

List of Tables

List of Figures

Abstract

The quantification of choice has been a major area of research for behavioural scientists for several decades. This is, in part, due to the discovery of the matching law (Herrnstein, 1961) which stipulates that relative response rates on concurrently available alternatives "match" the available relative reinforcement rates. This theoretical construct has been developed to describe response allocation in more complex situations, such as concurrent chains, and successfully describes both human and non-human behaviour. Typically, this phenomenon becomes evident when behaviour settles at an asymptote after several sessions of training where contingencies are held constant, and is often called "steady-state" behaviour. However, a fundamental question still remains: what causes matching – that is, what are the underlying momentary process(es) that produce matching? Researchers have suggested that what is necessary to answer this question is to take a molecular approach to the analysis of choice behaviour, thereby assessing choice in transition (Grace, 2002a). Recently, a new model of choice acquisition has been developed that appears to offer promise. It combines two separate mechanisms; a "winner-takes-all" categorical discrimination, and a linear-operator acquisition process (Grace & McLean, 2006). The initial results suggest this model could provide an alternative explanation for what underlies matching – that two separate processes are co-occurring in the acquisition of choice behaviour – allowing response allocation to be either linear or non-linear. This thesis extends the Grace and McLean model to include the situation of response strength 'carrying-over' from session to session to describe the process of acquisition gradually accumulating with experience. Moreover, additional assumptions have been added to describe temporal phenomena

and presumed discounting of previous experience on current choice behaviour. A steady-state version of the extended model was derived and, when fitted to published data sets, describes choice behaviour equally well when compared to existing models of steady-state choice. As a consequence of these additions, the Extended Decision Model (EDM) predicts a unique response allocation pattern – choice behaviour follows a bitonic function when initial-link durations were increased and the terminal-link delays were held constant. The results from experiments presented in this thesis support this prediction, whilst steady-state analyses found the EDM was parameter invariant – differences between parameters from two schedule types across several archival data sets were non-significant, while existing steady-state models had significant differences. These findings provide further support for the claim that the EDM and the Decision Model (DM) mechanisms provide unique and accurate descriptions of the molecular processes governing choice behaviour. Moreover, the implication from these results is that the underlying assumption of the EDM and DM – that choice is determined by the propensity to respond rather than conditioned reinforcement – appears to have further foundation. This challenges the assumptions of existing models of choice behaviour and presents the possibility that probabilistic approaches are perhaps more appropriate for describing response allocations than discrete estimates of relative value when contingencies change.

Chapter 1: Literature Review

The quantification of acquisition and steady-state choice behaviour has been of interest to researchers for numerous years. The experimental analysis of behaviour attempts to examine the fundamental laws that govern behaviour, and in an attempt to be specific and create testable hypotheses, researchers have created mathematical models to describe behaviour. The fundamental approach has been to describe relationships between behaviour and reinforcement (Thorndike, 1898). Perhaps the most influential formal representation of behaviour comes from steady-state concurrent-schedule research where concurrently available alternatives are repeatedly presented over several sessions. This methodology has been used in numerous steady-state experiments, leading to the discovery of an orderly relation between preference and reinforcement rates.

Steady-State

Concurrent Schedules.

Herrnstein (1961) designed an experiment where pigeons were confronted with two lighted keys in an experimental chamber. Responses to both keys were reinforced with access to food by independent variable-interval (VI) schedules. Pigeons were trained on a given pair of VI schedules where they received 60 reinforcers per session, lasting approximately 90 minutes, for a minimum of 16 sessions. Herrnstein found that the proportion of responses to each key was approximately equal to or "matched" the proportion of reinforcers obtained from that

key. This observation led to the development of the "Matching Law" that describes choice allocation:

$$\frac{P_1}{P_1 + P_2} = \frac{R_1}{R_1 + R_2} \qquad (1)$$

where P is behaviour (i.e., number of key pecks), and *R* is reinforcement (i.e., rate or number of food presentations). The alternatives are represented as subscripts *1* and *2*. The discovery of matching sparked a long and fruitful line of research. (Note: for the rest of the thesis, dependent and independent variables will be represented as uppercase letters and symbols except when a model uses the same letter or symbol for different variables then one variable will be lower case, whilst free parameters will be lower case italicised letters and symbols).

Baum and Rachlin (1969) examined the impact of relative reinforcement on pigeons' time allocation. Time spent standing on each side of the experimental chamber was recorded by switches in the floor, and reinforced by concurrent variable-interval schedules which provided grain through feeders located in either side of the chamber. They found the ratio of time spent on each side of the chamber was directly proportional to the ratio of reinforcements produced by standing on that side of the chamber, although there was a bias towards the right side of the chamber. Baum and Rachlin explicitly stated the matching law in terms of time allocation:

$$\frac{T_1}{T_1 + T_2} = \frac{R_1}{R_1 + R_2} \qquad (2)$$

where T is time allocation, *R* is reinforcement, and the subscripts *1* and *2* refer to the choice alternatives. Moreover, Baum and Rachlin introduced an important conceptual extension to the matching law, suggesting that the matching relationship might apply not to reinforcement rate *per se*, but to the relative "value" of the reinforcement associated with the alternatives. Baum and Rachlin suggested that independent variables such as the rate, amount, and immediacy of reinforcement, might combine to form a composite called "value". They used these variables as ratios to produce the concatenated matching law (CML):

$$\frac{T_1}{T_2} = \frac{R_1 A_1 I_1}{R_2 A_2 I_2} \qquad (3)$$

where, T_1 and T_2 are time spent on keys, R_1 and R_2 are rates of reinforcement, A_1 and A_2 are amounts of reinforcement, and I_1 and I_2 are the immediacies of reinforcement, associated with response alternatives *1* and *2* respectively. Baum and Rachlin suggested Equation 3 was the simplest way that different parameters of reinforcement could be combined to yield value.

These developments prompted researchers to examine the theoretical assumptions of the matching law. For example, Rachlin (1971) discussed whether the matching law was an empirical law subject to falsification, or a restatement of the experimental assumptions. He suggested that choice is a direct measure of reinforcement value and thus a tautology and it therefore cannot be disproved or

proved. Specifically, the matching law assumes that organisms choosing between alternatives are under no constraints except the contingencies of reinforcement in the experimental session, thus obtained reinforcement always equals reinforcement value. As these statements are logically the same, Rachlin suggests the matching law is merely a restatement of its own assumptions. However, Killeen (1972) argued that there was an important distinction between giving equations names, such as value, where there is equivalence between value and the experimental variable(s), and the matching of independent to dependent variables. He noted that a more general form of the matching relation would allow arbitrary transformations, while preserving the additivity assumption: log response rates are an additive function of relative reinforcer variables. Killeen suggested such a general form of the matching law could avoid the aforementioned tautology, provided that the parameters defining the functional transformations were invariant. Killeen's (1972) analysis was critical in the next step in the evolution of the matching law.

Baum (1974) introduced a generalized version of the matching law that included not only the strict matching of response to reinforcer allocation reported by Herrnstein (1961), but also two types of deviations from matching. Baum introduced the Generalized Matching Law (GML), expressed as follows:

$$\log\left(\frac{B_1}{B_2}\right) = a\log\left(\frac{R_1}{R_2}\right) + \log k \quad (4)$$

where B_1 and B_2 are response rates for alternatives *1* and *2*, and R_1 and R_2 are the reinforcer rates for alternatives *1* and *2*. Baum introduced the parameters *a* and log *k* to represent the sensitivity to reinforcer rate, and the bias, respectively. According to

this formula, the relationship between response allocation and reinforcer allocation is linear in logarithmic terms: a = slope of the response and reinforcer relationship and log k = y intercept in a linear relationship. Baum also defined two terms for deviations from matching: under-matching, which occurs when $a < 1$ where response allocation is less extreme than reinforcer allocation, and "over-matching", which occurs when $a > 1$ where response allocation is more extreme than reinforcer allocation. He also defined a term for response bias: an unknown but constant preference for one alternative over another, that is independent of the reinforcer ratio, represented by log k. Moreover, the GML reduces to strict matching when parameters a and log k are 1 and 0, respectively, and thus represents a *generalisation* of matching to a broader range of behavioural patterns. For example, Baum and Rachlin (1969) found the slope 1.0 indicated strict matching, but log k was positive, indicating a constant proportionality in response allocation favouring the right side of the chamber, which was unrelated to the reinforcer ratio. Baum (1974) analysed studies that found that response allocation was sometimes less extreme than reinforcer allocation, producing relative response slopes that are less than 1.0. He referred to this phenomenon as "under-matching" (Fantino, Squires, Delbruck, & Peterson, 1972).

Baum mentioned possible causes of under-matching, including poor discrimination between alternatives, insufficient changeover delay to prevent rapid switching between the alternatives, and deprivation level. He also suggested a few sources of bias: response bias (colour of keys, preference of movement, etc.), discrepancy between scheduled and obtained reinforcement, and the effects of qualitatively different reinforcers and schedules. In addition, Baum compared previous data from Fantino, Squires, Delbruck, and Peterson (1972) and La Bounty and Reynolds (1973) and found that if those authors had examined their data using a

bias parameter, their results would support matching. Therefore, the GML appears to describe choice behaviour across a range of response patterns – matching and systematic deviations from matching.

Concurrent Chains

Autor (1969) introduced concurrent chain schedules as a method of examining conditioned reinforcement. In a concurrent chains procedure, pigeons experience a chain of typically two schedules, the initial link and the terminal link. In this procedure, the initial link is presented as two identical keys that are illuminated white where responses that meet the schedule requirement for either key lead to mutually exclusive terminal-links. The key light then changes to another colour indicating the start of the terminal-link, which ends with food reinforcement. Autor arranged (Variable Interval) VI initial links and VI, DRO (Differential reinforcement for no responding), and VR (Variable Ratio) schedules during the terminal links, and found that relative initial-link response rates matched the relative terminal-link reinforcement rates regardless of the type of schedule. Because the terminal link response rates varied depending on the type of schedule, this result suggested that, regardless of schedule, initial-link response rates are a function of primary reinforcement during the terminal link, and not terminal-link response rates.

Some researchers have suggested that terminal-link stimuli acquire conditioned reinforcement value through Pavlovian conditioning (where reinforcement is not dependent on a response; Gallistel & Gibbon, 2000; Grace, 2002b). The idea is that contiguous 'pairings' of a terminal-link stimulus with primary reinforcement create an association of value to the stimulus, enabling the stimulus to reinforce a response similarly to a primary reinforcer (Fantino, 1977).

Evidence for this view was supported by research on the effect of stimuli eliciting instinctive behaviour using auto-shaping procedures (Brown & Jenkins, 1968). Brown and Jenkins presented an illuminated key to hungry pigeons that reliably predicted food. The pigeons continued to peck the key even though reinforcement was non-contingent on responding. Thus, the illumination of the key in close temporal proximity with food was sufficient to produce a key-pecking response. Jenkins and Moore (1973) found that different types of reinforcers changed the topography of the responses. For example, pigeons auto-shaped to peck on keys using food as a reinforcer pecked the keys with an open beak, closely emulating their eating responses to grains of wheat. Conversely, when pigeons were auto-shaped with water they pecked the keys with closed beaks and swallowing movements, similar to behaviour when drinking water. These results suggest that auto-shaped key pecking was caused by Pavlovian conditioning because 1) responses occurred even though reinforcement was non-contingent on key pecks, and 2) the responses to the keys were indicative of 'natural' or involuntary response to food, suggesting an association between conditioned stimuli and unconditioned stimuli rather than voluntary behaviour maintained by reinforcement.

A similar procedure to auto-shaping is used in the terminal link of concurrent-chain procedures. In concurrent chains, the completion of an initial-link schedule leads to onset of the terminal-link schedule, where stimuli indicate the following reinforcer that is non-contingent on responding. The assumption that relative value of the terminal-link stimuli determines preference is common to models of concurrent chains (Grace, 2002b).

Williams and Dunn (1991) examined the effect of conditioned reinforcement in the terminal link in a concurrent-chains procedure. They used VI 120-s initial links

and FI 20-s terminal links reinforced with food, and superimposed independent VI 30-s initial-link schedules using the same stimulus, but the scheduled ended without food. The absolute number of the no-food stimulus presentations was held constant while their relative frequency assigned to one or the other choice alternative was systematically varied. Their goal was to examine how preference was determined by the frequency of conditioned reinforcement. They found that, for all pigeons, preference was controlled by the relative frequency of terminal-link presentations, independently of the frequency of food presentations. This result is consistent with the view that initial-link response allocation depends on the relative frequency of the conditioned reinforcement provided by the terminal-link stimuli.

Thus, a possible conclusion from these experiments is that concurrent chains may be viewed as a concurrent schedule of conditioned reinforcement, with the value of the terminal-link stimuli determined by a Pavlovian process (Nevin, 1973). This supports the possibility of extending the matching law to concurrent chains, based on the assumption that the initial-link response allocation should match the relative conditioned reinforcement value of the terminal-link stimuli. However, later research suggested an interaction between initial-link and terminal-link durations on conditioned reinforcement, challenging the assumption of independence between concurrent-chain links and the extension of the matching law to concurrent chains.

Davison (1983) extended the CML to apply to concurrent chains by including separate parameters for initial-link and terminal-link schedules. He suggested that a concurrent-chains matching model should reduce to the generalized matching law when the terminal-link delay was zero. Davison proposed that in that case, his model might provide an account of results from both concurrent schedules and concurrent chains. He suggested the following equation:

$$\left(\frac{B_L}{B_R}\right) = b\left(\frac{U_{1R}}{U_{1L}}\right)^{a1}\left(\frac{U_{2R}}{U_{2L}}\right)^{a2} \quad (5)$$

where B_L is responses on the left key, B_R is responses on the right key, *b* is a bias term, U_{1L} and U_{1R} are the mean initial-link intervals between conditioned reinforcers, U_{2L} and U_{2R} are the mean terminal-link delays to primary reinforcement, and *a1* and *a2* are sensitivities to initial-link conditioned reinforcement and terminal-link primary reinforcement rates, respectively.

Research using this model has found mixed results. Davison (1983) using a 3-second changeover delay (COD) and various initial- and terminal-link schedules, found that sensitivity to initial-link duration was independent of both initial- and terminal-link durations. However, other researchers found systematic variation in initial-link sensitivity, challenging the matching assumption of parameter invariance, when initial-link durations change (Alsop & Davison, 1988; Fantino & Davison, 1983).

A fundamental problem with the CML model when applied to concurrent chains was that it failed to capture the effect of changing initial and terminal-link durations on sensitivity to terminal-link reinforcement variables. Davison (1987) hypothesised that terminal-link sensitivity might change with changing initial-link delay, which challenges the matching assumption of parameter invariance and suggests an interaction between initial- and terminal-link delays on terminal-link sensitivity. This result has been supported in a number of studies (Fantino & Davison, 1983; MacEwen, 1972; Williams & Fantino, 1978). Davison (1987) concluded that all

models performed equally well (or poorly). Moreover, other theories of choice that use different mechanisms for determining choice behaviour appeared better able to describe temporal context effects.

Fantino (1969) presented a model for describing choice based on the degree of reduction in the expected time to primary reinforcement signalled by the terminal-link stimuli. According to Fantino's Delay-Reduction Theory (DRT), the strength of a terminal-link stimulus as a conditioned reinforcer depends on the reduction in delay to reinforcement signalled by the terminal link, relative to the average time to reinforcement from the onset of the initial links. Choice is then predicted as the relative delay reduction of the terminal-link stimuli:

$$\frac{R_L}{R_L + R_R} = \frac{T - t_L}{(T - t_L) + (T - t_R)} \qquad (6)$$

where R is responses, T is the expected time to reinforcement from the onset of the initial links, and t is the expected time to reinforcement from the onset of each terminal link, where the subscripted L and R are the left and right alternatives, respectively. This equation only applies when both $t_L<T$ and $t_R<T$ (i.e., both terminal links signal a reduction in time to reinforcement). If one alternative does not signal a reduction in delay to reinforcement, Fantino noted that the model should predict exclusive preference for the other (i.e., $t_L<T$, and $T_R>T$ preference will be exclusive for the left alternative, i.e. 1, and when $t_L>T$, and $T_R<T$ preference will be exclusive for the right alternative, i.e. zero).

Fantino (1969) showed that preference for a VI 30-s compared to VI 90 s terminal-link decreased as the initial links were changed from VI 30 to VI 600 s. This

result is predicted by Equation 6, and is known as the 'initial link effect'. By contrast, the matching law was unable to describe this effect. Fantino concluded that the relative delay reduction, rather than relative reinforcement rates, appears to determine choice allocation.

Squires and Fantino (1971) identified an unlikely prediction of DRT: that preference in the initial links should be indifferent when terminal-link delays are equal, irrespective of the relative values of the initial-link schedules. Thus Equation 6 predicts that differences in primary and conditioned reinforcement rates have no effect on preference, provided that the relative delay reduction is equal. They tested this assumption using a concurrent-chains procedure in which the initial-link schedules were varied while the terminal links were constant. Squires and Fantino found that response allocation favoured the initial link that provided the more frequent access to its terminal link. They proposed an extension of DRT in which the overall rate of primary reinforcement on each key, r, was added as a new factor in the model:

$$\frac{R_L}{R_L + R_R} = \frac{r_L\ (T - t_{2L})}{r_L\ (T - t_{2L}) + r_R\ (T - t_{2R})} \quad (7)$$

where R is response rate, T is the expected time to primary reinforcement from the onset of the initial links, t_2 is the average duration of terminal-link delay and L and R represent the left and right alternatives. The rate of primary reinforcement, r, is calculated (for the left alternative) as $r_L = n_L/(t_{1L} + n_L t_{2L})$, where n_L is the number of primary reinforcements obtained per terminal-link entry, t_1 is the initial-link delay and t_2 is the terminal-link delay.

Squires and Fantino found that their extended model provided a better account of their results than either the matching law or the earlier version of delay-reduction theory. In addition to describing the initial-link effect, it was able to predict matching of choice responding to the relative rates of reinforcement when alternatives were equal. These results appeared to challenge the assumption of the matching law that response allocation should match the value of the terminal-link stimuli as conditioned reinforcers, and suggested an interaction of sensitivities rather than invariance.

Grace (1994) proposed a model for concurrent chains based on the CML that attempted to resolve the difficulties with parameter invariance that Davison (1987) had identified. He began by re-analysing Fantino and Davison's (1983) data, plotting the sensitivity to relative conditioned reinforcement rates (i.e., relative initial-link schedules) against the average time spent in the initial link. He found no consistent trend and suggested that as only one study had previously found that sensitivity had varied with initial-link duration (Alsop & Davison, 1988), whereas two other studies found no consistent trend (Davison, 1983; Fantino & Davison, 1983), he assumed that initial-link sensitivity was invariant, although adding a caveat that this assumption may need revising in light of new data.

Next, after examining the terminal-link sensitivity, he concluded that terminal-link sensitivity is affected by temporal context, specifically the overall duration of the initial and terminal links (Fantino & Davison, 1983; MacEwen, 1972; Williams & Fantino, 1978). Specifically, Grace suggested that sensitivity to terminal-link reinforcement variables was a function of the ratio of the average times spent in the terminal and initial links. This assumption served as the basis of his contextual choice model (CCM).

$$\left(\frac{B_L}{B_R}\right) = b\left(\frac{U_{1r}}{U_{1L}}\right)^{a1}\left[\left(\frac{U_{2r}}{U_{2L}}\right)^{a2}\left(\frac{X_{IL}}{X_{IR}}\right)^{ai}\right]^{\left(\frac{T_t}{T_I}\right)^k} \quad (8)$$

Equation 8 is based on the CML (Equation 5), but includes two additional features: Temporal context is represented by $(T_t/T_I)^k$, where T_t and T_I are the average time spent in the terminal and initial links per cycle, respectively. Because of the laws of exponents, the effective sensitivity to terminal-link reinforcer rate is $a2*(T_t/T_I)^k$, and thus Equation 8 predicts that sensitivity will increase as terminal-link duration increases with respect to initial-link duration, and decrease as initial-link duration increases with respect to terminal-link duration. These are the effects of temporal context on sensitivity that Davison (1987) noted were problematic for the CML. Second, the ratio $(X_{IL}/X_{IR})^{ai}$ represents an additional reinforcement variable that differentiates the terminal links, for example reinforcer magnitude. The *k* parameter is a scaling exponent, and determines how much terminal-link sensitivities change as a function of temporal context. Grace (1994) found that *k* was only necessary (and generally took values < 1) when the terminal links were not differentially signalled, and that otherwise was set equal to 1 (and thus drops out of the calculation).

Several other features of CCM are worth noting. First, CCM reduces to the GML when the terminal link duration is zero, as raising the terminal-link ratios (in brackets in Equation 8) to an exponent of zero means that those ratios equal 1 and thus do not contribute to preference. Second, CCM assumes that terminal value is determined by the variables in brackets in Equation 8 (i.e., average time to primary

reinforcement and reinforcer magnitude), but is modulated by the temporal context. Thus, CCM makes a separation between conditioned reinforcement value and the behavioural expression of relative value as preference, analogous to the traditional distinction between learning and performance. The learning of the value of an alternative occurs in the terminal link, which is based on the conditioned reinforcement of each alternative, whilst expression of relative value occurs in the initial-link where preference is measured as the relative response rates to the alternatives.

Grace fitted his model to all the available concurrent-chains research that had a minimum of four data points for each subject. He analysed studies using the following criteria: preference changed at least 12.5% from baseline, alternatives had the same type of terminal-link schedule (either both VI or both FI), and terminal-link magnitudes were equal for each alternative. This yielded a total of 92 individual-subject data sets from 19 studies. Across studies, CCM accounted for approximately 90% of the variance in initial-link response allocation and described response allocation better than existing models such as delay-reduction theory (DRT) and incentive theory (Killeen, 1982; Squires & Fantino, 1971). In terms of variance accounted for, CCM provided a description of the data that was comparable to that given by the generalized matching law for concurrent-schedules responding (Baum, 1979). However, CCM could only describe half of the total number of studies according to criteria suggested by Davison (1987). These were that 1) the slope and intercept estimates from a regression of obtained on predicted data should differ from 1.0 and 0.0 by less than two standard deviations, 2) the standard error of the prediction should be 0.10 or less, and 3) the standard deviations of slope estimates should be within one tenth of the slope estimates. Grace found the majority of the data

sets met the first criterion (89 out of 92) but questioned whether the second criterion was appropriate for the significant number of data sets in his analysis that had relatively few data points. When he analyzed the data sets that had greater than 12 data points per set, 30 out of 32 met the second criterion, but when data sets that had 12 or fewer data points per set were analyzed only 23 out of 60 met the second criterion.

Mazur (2001) proposed an alternative model for concurrent chains choice behaviour, based on previous work that had identified that the following hyperbolic function described the value of delayed reinforcers (Mazur, 1984):

$$V = \frac{A}{(1+kD)} \qquad (9)$$

equation 9, known as the hyperbolic decay model, where V is the value of a reinforcer of amount A delivered after delay, D. The parameter k is the rate at which the discounting occurs. Equation 9 has proven to be a useful model for temporal discounting – how reinforcement value decreases as a function of delay, and has been widely applied to human choice (see Green & Myerson, 2004, for review). To develop Equation 9 into a model for concurrent chains, Mazur (2001) calculated the value of a terminal link with a more general version of Equation 9 that can apply to a distribution of delays:

$$V = \sum_{I=1}^{N} P_I \frac{A}{(1 + kD_I)} \qquad (10)$$

where p_I is the probability of a reinforcer delay of D_I seconds. This equation implies that the total value of an alternative with variable delays is the weighted average of all delays, with their values decaying hyperbolically (Equation 9), given their probability of occurrence. Therefore, this method is able to describe both fixed and variable delays, making a comparison between fixed and variable schedules of reinforcement possible. Mazur then extended his model to concurrent chains by making the following assumptions: a) initial-link choice matches the relative rates of conditioned reinforcement, as described by the generalized matching law; b) the value of each terminal link depends on the time from the onset of that link to primary reinforcement, c) the value of the initial links depends on the time from the onset of that link to primary reinforcement, and d) response allocation is based on the relative value added at the onset of a terminal link (the increase in value when the terminal link is entered). These assumptions are formally expressed as follows:

$$\left(\frac{B_L}{B_R}\right) = b\left(\frac{R_{I1}}{R_{I2}}\right)^{aI}\left(\frac{V_{T1} - a_T V_I}{V_{T2} - a_T V_I}\right) \quad (11)$$

where B_L and B_R are responses made to the left and right alternatives respectively, b is bias, R_{I1} and R_{I2} are initial link reinforcement rates for alternatives 1 and 2, and a_I is the sensitivity parameter to the initial links. In the right hand bracket is the unique feature of Mazur's model, the calculation of value addition. V_{T1} and V_{T2} are the values associated with the onset of terminal links 1 and 2, and V_I is the value for the onset of the initial links and a_T is the sensitivity parameter to the terminal links.

This model is similar to a model developed for concurrent chains using hyperbolic decay (Davison & Smith, 1986). Davison and Smith suggested that reinforcer value depended on a summation of the delays from initial- and terminal-links. This is comparable to the addition of value to an alternative that is assumed by HVA to occur when subjects enter the terminal-link. Mazur compared HVA with CCM and DRT using the same data sets as Grace (1994) had analyzed. Mazur found all three models described choice about equally well. However, Mazur identified an important difference between the models: both CCM and HVA are based on the generalized matching law, and thus assume that initial-link response allocation matches the relative rate of conditioned reinforcement. As the terminal links are signalled by distinctive stimuli, concurrent chains may be viewed as an extension of concurrent schedules in which conditioned, rather than primary, reinforcement controls choice. In contrast, DRT measures the total time to primary reinforcement and assumes no direct role for rate of conditioned reinforcement but instead assumes relative rate of primary reinforcement is responsible for the effects of unequal initial-link schedules on preference (Fantino & Romanowich, 2007).

Temporal Context and Conditioned Reinforcement: Concurrent-Chains

One important difference between CCM and DRT is whether the value of a conditioned reinforcer depends on temporal context. According to DRT, temporal context determines value directly, whereas CCM assumes that value is independent of temporal context. Grace and Savastano (2000) examined this issue, using a transfer design with occasional choice probe trials which tested the ability of terminal-link stimuli to evoke responding and reinforce responding in a new initial-link context.

They compared predictions based on whether value was independent (CCM) or dependent (DRT) on temporal context. Grace and Savastano found that subjects learned the temporal relations between stimuli and reinforcement independently of temporal context – response allocations in the probe tests were determined by the reinforcer delay signalled by the terminal links independently of duration of the initial links or the other terminal link in baseline (but cf. O'Daly, Meyer, & Fantino, 2005). Grace and Savastano suggested that their results were consistent with the view that terminal-link value was predicted by the stimulus-reinforcer relation, that is, by a Pavlovian process.

Grace (2004) used a three component procedure where each component contained three separate concurrent chains in which terminal-link variable-interval schedules were a constant ratio but their average duration increased across components. Grace used the different terminal-link schedules to assess variations of the Contextual Choice Model (Grace; 1994, CCM), Delay Reduction Theory (Fantino; 1969, DRT), and the Hyperbolic Value Added model (Mazur; 2001, HVA). He found that each model described approximately 90 - 93% of the variance in the obtained data. However, the relationship between preference and T_t/T_I suggested a negatively accelerating curve. This result challenges the DRT assumption of a positive accelerating function and implies an asymptotic level of changing the ratio of terminal- and initial-link durations. The best fitting model was CCM using a $2T_t/(T_t+T_I)$ context scaling parameter. In addition, results supported the assumption of CCM and DRT, that response allocation should remain constant when the initial- and terminal-link duration was increased by the same factor, suggesting ratio invariance. Grace then extended the relative range of terminal- to initial-link duration. He found CCM and HVA both described approximately 93% of the variance in the obtained

data, but there was no systematic increase in preference as the ratio of T_t/T_I increased. However, Grace found that CCM, with the $2T_t/(T_t+T_I)$ contextual scaling parameter, was the only model with no statistically significant residuals (the difference between obtained and predicted scores). Grace suggested that the CCM $2T_t/(T_t+T_I)$ contextual scaling parameter may be analogous to the ratio of average inter-reinforcer intervals to signalled delay of reinforcement in classical conditioning *(C/T),* which Gibbon and Balsam (1981) found determined the rate of acquisition in autoshaping. The similarity of temporal context effects in concurrent chains and autoshaping is also consistent with the view that a common underlying process – Pavlovian conditioning – determines response allocation in concurrent chains and the acquisition of autoshaping in discrete trial procedures. This supports the extension of the matching law to concurrent chains – that concurrent chains can be viewed as a concurrent schedule of conditioned reinforcement, and that initial-link responding matches the relative frequency and value of the terminal-link stimuli according to the CML (Autor, 1969; Herrnstein, 1964).

Temporal Context: Initial-Link and Terminal-Link effects

Temporal context – the overall duration of the initial and terminal links – has significant effects on response allocation in concurrent chains. These results have challenged the matching law and have been incorporated into arguably the most successful version of concurrent-chains behaviour – CCM (Grace, 1994). The two primary phenomena are the initial-link effect (Fantino, 1969) and the terminal-link effect (MacEwen, 1972). The initial-link effect has been previously discussed and has been consistently replicated (Grace, Berg, & Kyonka, 2006; Mazur, 2004). The

terminal-link effect is the observation that preference for the alternative associated with a shorter delay to reinforcement becomes more extreme as the absolute duration increases whilst the ratio of the delays are held constant (MacEwen, 1972). This effect has been obtained with both VI and FI terminal links (MacEwen, 1972; Williams & Fantino, 1978), equal-duration terminal links that vary in terms of reinforcer magnitude (Navarick & Fantino, 1976), and probability of reinforcement (Spetch & Dunn, 1987). Thus, the well-replicated existence of these phenomena implies that theories of concurrent chains that presume to be a comprehensive description of choice need to describe the effects of both links.

The conclusion from the aforementioned studies is that behaviour appears to be lawful under different choice situations, supporting the possibility of a general model of choice in concurrent chains, most probably mediated via conditioned reinforcement. However, applications of both the original and concatenated matching law to concurrent chains by Herrnstein (1964) and Davison (1983), respectively, have been complicated by the presence of temporal context effects (although Wardlaw and Davison (1974) were able to predict an initial-link effect using a semi-contextual model). These phenomena have informed the development of models of steady-state choice, such as DRT, CCM, and HVA, which are based on the matching law but make additional assumptions. Nevertheless, an important limitation of these models is that they apply only to response allocation at the molar level and are silent about the molecular processes active in momentary behaviour – the acquisition of choice.

Choice - Acquisition

Acquisition is the process where novel behaviour becomes learned as a consistent part of an organism's repertoire. The initial research of Thorndike (1898) and Pavlov (1927) was seminal in the development of comprehensive theories of conditioning. The quantification of these ideas soon followed and led to the development of models based on a *linear operator* approach – that is, in which a theoretical construct representing learning (such as associative value or response strength) was updated on a trial-by-trial basis by a constant proportion of an asymptotic or final value (Bush & Mosteller, 1951; Estes, 1950; Rescorla & Wagner, 1972). This approach can provide an account of choice, if it is assumed that choice behaviour is the direct result of the comparison of the relative values of the available alternatives (Skinner, 1986). Thus, many researchers have assumed that an adequate theory of the acquisition of choice could be gained through understanding the change in response strengths of the individual alternatives through a linear-operator process (Williams, 1994). Variations using this approach have been used to develop various choice acquisition models.

Concurrent Schedule Models

Mazur (1992) examined choice behaviour in transition with variable ratio (VR) and variable interval schedules. In Experiment 1, response rates were examined in 50 periods of transition with two VR schedules. Each condition began with the same reinforcer probability (i.e., VR value) for each alternative, and then switched to unequal probabilities. Mazur found when the ratio of probabilities was constant there was faster acquisition for higher probabilities than when the probabilities were low. In

Experiment 2, conditions began with equal VI schedules of reinforcement, then one alternative's schedule changed while the overall reinforcement rate was held constant. Mazur found that the speed of acquisition was approximately the same regardless of the relative reinforcement rate. Mazur presented a simple linear-operator model of acquisition that predicts response-by-response changes in the probability of choosing each of the two alternatives to describe his data. The model includes an equation for reinforcement and non-reinforcement. The equations are as follows:

$$\Delta V_I = r(1 - V_I) \qquad (12)$$

and

$$\Delta V_I = n(-V_I) \qquad (13)$$

where i is each response, V_I is the strength of response $_I$, ΔV_I is the change in strength of V_I, and r and n are the learning-rate parameters for reinforcement and non-reinforcement respectively. Mazur restricted the range of values for V_I, r and n from 0 to 1. Moreover, he assumed a matching-type rule to calculate the choice probability for each alternative:

$$P_1 = \frac{V_1}{V_1 + V_2} \qquad (14)$$

where P_1 is the probability that response 1 will occur (note that $P_2 = 1- P_1$), and V_1 and V_2 are response strengths for alternative 1 and 2 respectively. Mazur then tried a number of different parameter values for *r* and *n* which produced similar acquisition rates in Experiment 1. Mazur found that his linear-operator model provided a better account of the data than either Bush and Mosteller's linear-operator model (1955) or the kinetic model (Myerson & Miezin, 1980) did. He then used parameter values found in Experiment 1 to produce predictions for Experiment 2. Mazur found his linear-operator model was able to describe the acquisition curves for each of the conditions. He also found his model successfully described response allocation in a similar experiment (Mazur & Ratti, 1991). Mazur concluded that these results suggested that; 1) the strength of an alternative increases with reinforcement and decreases with non-reinforcement; and 2) a linear-operator process can describe the change in strength.

Mazur (1996) examined choice behaviour between two keys using a single VI 30s schedule that was assigned on each cycle to one key. The assignment followed a quasi-random sequence where once a key leading to reinforcement was assigned no further reinforcement was available until a reinforcer was delivered. The procedure also included a 2s change over delay. Experiment 1 varied the probability of reinforcement on key 1; 50% was delivered for several sessions, then either 70% or 90% for one, two, or three sessions, and then 50% for the remaining sessions. The subjects' choice proportions post-transition exhibited spontaneous recovery, or a return to similar levels of responding as in earlier sessions. In Experiment 2, a three-day rest period was added before a transition, and spontaneous recovery was quicker than when no rest period occurred. In an attempt to describe the acquisition process and the change in behaviour post-transition, Mazur used a temporal weighting rule

based on the previous five sessions as a modification to his earlier linear-operator model (Devenport & Devenport, 1994; Mazur, 1992). The resulting composite model describes both processes, leading Mazur to conclude that the influence of prior reinforcement diminished with time, and that a weighted average of response strengths over the past several sessions was an accurate mechanism for describing the behaviour at the start of a session.

Mazur (1997) examined acquisition using a single variable-interval schedule for a two-key choice procedure with probabilistic reinforcement. He attempted to assess how the overall rate of reinforcement affected the acquisition of preference for the alternative with the higher reinforcement rate. He found that response allocation reached similar asymptotic levels regardless of overall reinforcer rate, but was faster at higher overall rates. Moreover, the changes in the rates of transition were considerably smaller than the changes in the rates of reinforcement. Mazur compared the probability of responses after reinforcement with non-reinforcement and found pigeons tended to make more responses to the same alternative after reinforcement than after non-reinforcement.

In addition, Mazur examined the effect of the speed of changing contingencies on preference. He found acquisition rates were higher when contingencies changed every 1 or 2 sessions than when they changed on average every 8 sessions. He also found using 8 session transitions, subsequent acquisition rates were much lower than in the first series. Mazur compared the obtained data with his linear-operator model of choice acquisition (Mazur, 1992) and found that overall reinforcement had an effect on acquisition that was inconsistent with the predictions of this model. In addition, changes in response rates during transition were much less than the change in reinforcement, contrary to the predictions of the kinetic model (Myerson & Miezin,

1980). Mazur concluded that the effectiveness of individual reinforcers changes depending on the overall context of reinforcement.

Davison and Baum (2000) examined acquisition rate in multiple-component concurrent schedules in which the reinforcement contingencies were frequently changed. Pigeons were exposed to a mixed concurrent VI VI schedule with seven components. The conditions lasted 50 sessions. Each of seven components was associated with a different reinforcer ratio. There were two parts to the experiment. In part one, the number of reinforcers varied from 4 to 12 per component with an overall reinforcer rate of 2.22 reinforcers per minute, and in part two, the number of reinforcers varied across the same range as part one, but with 6 reinforcers per minute.

Davison and Baum found that individual reinforcers had an effect on response allocation, and that acquisition was faster with higher overall reinforcement rates. When the effects of carry-over of response allocation were excluded, the number of reinforcers appeared to make no difference in the speed of behaviour change in a new component. Within components, successive reinforcers from the same alternative had a smaller effect than the previous reinforcer (that is, diminishing returns), whereas a single reinforcer from the opposite alternative always had a large effect ('disconfirmations'). Carry-over effects lasted until up to five or six reinforcers into the next component. To describe these effects, Davison and Baum presented a "local" model for response strength on each alternative based on Davison and Jenkins' (1985) response-reinforcer contingency discriminability model. Davison and Baum's model assumed three processes; 1) Each reinforcer is allocated to one of two response-related accumulations, 2) Later events have less strength for controlling behaviour, and 3) During blackouts the accumulations of reinforcers to alternatives become less differentiated - accumulations should "leak" into each other. For the $_i$th reinforcer

delivery, which may be on the left (R_{LI}) or right alternative ($R_{r,I}$), the equations for the accumulation differentiation process are:

$$R_{LI} = R_{L,I-1} + p_d R_{LI} + (1 - p_d) R_{r,I} \quad (15)$$

and

$$R_{rI} = R_{r,I-1} + p_d R_{rI} + (1 - p_d) R_{L,I} \quad (16)$$

where either $R_{L,I} = 1$ and $R_{r,I} = 0$, or vice versa. R is the reinforcer accumulation prior to (subscript $_{I-1}$) and after (subscript $_I$) reinforcer $_I$. p_d is the response-reinforcer discriminability parameter (for perfect discriminability, $p_d = 1.0$). The equations for the loss of reinforcers and loss of reinforcer differentials are:

$$R_{L,T} = p_\Delta p_e R_{L,I-1} + (1 - p_\Delta) p_e R_{r,T-1} \quad (17)$$

and

$$R_{r,T} = p_\Delta p_e R_{r,T-1} + (1 - p_\Delta) p_e R_{L,T-1} \quad (18)$$

where R is the reinforcer accumulation prior to (subscript $_{T-1}$) and after (subscript $_T$) the end of the time period. The values of parameters p_Δ and p_e depend on the time period chosen, p_Δ is accumulator discriminability at the end of a time period, while p_e is arranged extraneous reinforcer discriminability. Both equations are included in the entire model. Davison and Baum found their model provided a good description of how response allocation changed within components. They also suggested that the

model could predict typical steady-state results (i.e., matching and under-matching) with appropriate p_d values.

The assumption of a time period accumulation suggests that although Davison and Baum's model can describe changes in allocation on a response-by-response basis, there appears to be an implicit assumption of temporal effects. The authors suggest their model can also describe the effects of forgetting and hysteresis after a change in experimental conditions (Davison & Hunter, 1979). Moreover, they suggest their model can also predict nonexclusive preference on concurrent VI extinction (Davison & Jones, 1998). Davison and Baum suggest that their model might be a viable alternative to molar (Rachlin, Green, Kagel, & Battalio, 1976), molecular maximization (Silverberg & Zirax, 1982) and melioration models (Herrnstein & Vaughan, 1980).

This attempt at describing the process that underlies matching at steady-state has produced numerous models of choice. However, the evidence for many behavioural theories appears mixed (Bailey & Mazur, 1990; Herrnstein & Heyman, 1979; Horner & Staddon, 1987), and research comparing models of the acquisition of choice in concurrent schedules has found difficulties for most accounts (Mazur, 1997; Staddon, 1988). However, certain phenomena are obtained consistently, such as hysteresis effects (Bailey & Mazur, 1990; Davison & Baum, 2000), and faster acquisition with higher rates of reinforcement and probability (Bailey & Mazur, 1990; Mazur & Ratti, 1991; Mazur, 1992, 1997). Nevertheless, linear-operator models have often been used to model acquisition of choice in concurrent schedules (Dreyfus, 1991; Lea & Dow, 1984; Mazur, 1992; cf. Mazur, 1997). More recently, the investigation of choice acquisition has been extended to concurrent-chains procedures (Grace, 2002a; Mazur, Blake, & McManus, 2001).

Concurrent-Chains Acquisition

Mazur, Blake, and McManus (2001) used the concurrent-chain procedure to assess how choice behaviour changed when delays or probabilities of reinforcement are switched. Mazur et al. assumed that pigeons would take longer to discriminate a change in probabilistic schedules than changes in delays, because continued exposure is needed to discriminate a change in probabilistic schedules. However, Mazur et al. found slower change in preference when delays were switched than when probabilities were switched. Analysing the data, reinforcer-by-reinforcer, Mazur et al. found that single reinforcers had a greater impact on initial-link response allocation in the probabilistic than delay conditions. Moreover, response rates to the rich alternative were higher after the presentation of the rich-key reinforcer than a lean reinforcer, and this effect persisted for at least 20 responses after a reinforcer. In addition, the effect of individual reinforcers after a switch persisted in the probabilistic conditions, but rapidly diminished in the delay conditions. Mazur et al. concluded that slow transitions during variable interval (VI) and variable ratio (VR) schedules are not due to inherent difficulties in discriminating change in variable schedules, but rather are determined by a combination of current and prior reinforcement.

Grace (2002b) attempted to examine the common assumption of existing molar models of choice in concurrent chains, such as DRT, CCM and HVA, that relative terminal-link value determines initial-link preference. He described this assumption as the "value hypothesis". Grace examined the value hypothesis in two acquisition experiments in concurrent chains. In Experiment 1, he compared the acquisition of initial- and terminal-link responding after the terminal-link schedules were reversed.

Temporal control during the terminal links was assessed with occasional no-food trials lasting 60s ending without reinforcement, similar to the peak procedure (Roberts, 1981). Grace took the responding on the no-food peak trials as a measure of the terminal-link value – response rates in these trials were non-reinforced, and thus assumed to be maintained by the conditioned value of the terminal stimuli (Grace & Nevin, 1999). Using a successive reversal design, he reversed the terminal-link delays while maintaining the initial-link keys leading to the terminal-link stimuli. Grace suggested the value hypothesis would predict that changes in response allocation on no-food peak trials should occur prior to, or concurrently with, shifts in preference, as preference is directly related to terminal-link stimuli. This prediction was supported – acquisition of temporal control in the no-food terminal link trials occurred more rapidly than changes in initial-link response allocation.

In Experiment 2, Grace assessed the speed of acquisition when the terminal-link schedules were reversed and the terminal-link stimulus-reinforcer relations were changed, compared to when the stimulus-reinforcer relations were maintained. Grace assumed the value hypothesis would predict faster acquisition if the terminal-link stimuli continue to signal the same reinforcement schedules than if they signal different schedules. He found preference changed more rapidly when terminal-link stimulus-reinforcer relations were maintained than when they were changed. Thus, in both experiments, the results supported the value hypothesis – initial-link response rates were mediated by terminal-link stimuli-reinforcer contingencies. However, an earlier study by Grace and Nevin (1999) had questioned the value hypothesis in terms of an association between choice and timing.

Grace and Nevin (1999) investigated the role of timing processes in choice by combining the peak procedure (occasional trials where terminal-link stimuli were

presented without food reinforcement; Roberts, 1981) with concurrent chains. Pigeons were first trained in a peak procedure where delays to reinforcement were differentially signalled. After 25 sessions, the peak procedure was "embedded" within a concurrent-chains procedure, so some terminal-link stimuli presentations continued past the time of reinforcement on food trials, so that temporal control could be assessed. The inter-trial interval in the peak procedure was replaced by initial-link choice stimuli (white keys), while the terminal-link stimuli were the same as the stimuli in the peak procedure. Grace and Nevin found that the temporal control on the no-food trials in the peak procedure was the same as in the peak procedure embedded into the concurrent-chains design – so inclusion of the choice phase did not disrupt temporal control. The pigeons were then reintroduced to the peak procedure alone and then the combined concurrent-chains and peak procedure, this time substituting a larger delay for the smaller alternative and thereby reversing expected preference. However, choice allocation did not immediately reflect the changed delay when the pigeons were returned to concurrent chains, despite the fact that terminal-link responding was unchanged and showed that the pigeons were accurately timing the changed delays. Grace and Nevin suggested this result challenges theories that suggest choice is based on a fundamental timing process, such as scalar expectancy theory (Gibbon, Church, Fairhurst & Kacelnik, 1988). Moreover, this result is also problematic for the value hypothesis, which would predict that initial-link response allocation should reflect the change in terminal-link value, especially as pigeons were able to accurately time the change in delay to reinforcement in the same procedure. Grace and Nevin's (1999) results suggest that the factors determining choice in transition may be different from those responsible for choice at the molar level, and raise questions about the status of 'value' as an explanatory construct for choice.

Models for Acquisition in Concurrent Chains

Grace (2002a) tested two models of acquisition of preference in a concurrent-chains procedure. Pigeons were trained with pairs of fixed interval and variable interval terminal-link schedules in a successive reversal design (schedules reverse after conditions lasting 20 sessions each). He compared predictions from a generic linear-operator model (LINOP) with a generic memory representation of delays model (MEMREP). Grace made a number of assumptions to derive the LINOP model, including that initial-link preference was determined by the relative value of terminal-link stimuli as conditioned reinforcers, and the asymptotic value of a terminal link was determined as a function of the distribution of delays to reinforcement signalled by terminal-link onset. Moreover, Grace assumed that a linear-operator rule described acquisition: the change in terminal-link value on each trial was a constant proportion of the difference between the current value and the asymptotic value for that reinforcement schedule. This is expressed as:

$$\Delta \mathrm{V}_{\mathrm{N}+1} = r\left(V_{asymp} - \mathrm{V}_{\mathrm{n}}\right) \quad (19)$$

where r is a learning rate parameter, V_{N} is the 'value' at trial $_{\mathrm{N}}$, and V_{asymp} is the asymptotic value for the terminal-link schedule. Grace showed that the LINOP model predicted that the rate of acquisition of preference (using a normalised measure of percentage of eventual asymptotic change, to compare acquisition across conditions with different asymptotic levels of preference) depended on the values of the schedules that preceded and followed the reversal: the rate of acquisition was faster when pigeons were exposed to relatively low-valued schedules prior to reversal (e.g.,

FI not VI, or long duration not short duration), and relatively high-valued schedules after the reversal (e.g., VI not FI, or short duration not long duration). Thus, the fundamental prediction of the LINOP model is that relatively higher-valued schedules have stronger effects over response rates: slowing acquisition when prior to the transition, and speeding acquisition if presented post-reversal.

Grace compared LINOP with another generic model, the Memory-Representational Model (MEMREP). The basic assumption of the MEMREP model is that pigeons represent reinforcement delays as an internal memory, and sample delays from their memory when responding during the initial links, choosing the alternative with the shorter remembered delay. This assumption is similar to scalar expectancy theory (Gibbon et al., 1988). The basic equation is:

$$\frac{B_L}{B_L + B_R} = f[P(X_L < X_R)] \qquad (20)$$

where X_L and X_R are delays sampled from the left and right memories, respectively. Grace assumes f is an identity function. Grace (2002a, Appendix B) derived predictions for the MEMREP model assuming that each memory contained 12 values (at asymptote, each interval has 1/12 probability of being sampled on any given comparison). However, during acquisition, the probabilities of the preceding schedule intervals decrease from 1/12 to 0, while the probabilities of the following schedule increase from 0 to 1/12. This models the effect of changing memories as the pigeons are exposed to the new schedules. MEMREP assumes that pigeons sample delays from their memories and then compare them in terms of their value, i.e. both immediate delays are preferred over less immediate delays and thus MEMREP is

similar to models such as scalar expectancy (Gibbon, Church, Fairhurst, & Kacelnik, 1988) and rate expectancy theory (Gallistel & Gibbon, 2000).

Grace arranged different pairs of fixed-interval (FI) and variable-interval terminal-link schedules in a successive-reversal design. There were four different schedule pairs: VI 8 s VI 16 s (Variable Interval Short duration: VS), VI 16 s VI 32 s (Variable Interval Long duration: VL), FI 8 s FI 16 s (Fixed Interval Short duration: FS), and FI 16 s FI 32 s (Fixed Interval Long duration: FL). He studied all possible transition pairs between these schedules (that is, 16 transitions: VSVS, VSVL, VSFS, VSFL, VLVS, VLVL, VLFS, VLFL, FSVS, FSVL, FSFS, FSFL, FLVS, FLVL, FLFS, FLFL). The key questions were whether there were effects of schedule type (e.g., FI vs VI) and duration (short vs long) on the acquisition of preference.

The important difference between models is that MEMREP is based on a "winner takes all" response rule and does not differentiate between the magnitudes of difference between alternatives, whereas LINOP predicts increments of response strength depending on the size of the difference between current response strength and the asymptote. Despite these differences, Grace showed that both models predicted that acquisition should be faster when FI schedules preceded a reversal, and when VI schedules followed a reversal. Both predictions were confirmed. However, the models differed in terms of their predictions for schedule duration: MEMREP predicted that acquisition should not depend on duration, but LINOP predicted faster acquisition for long schedules preceding a reversal, and for short schedules following the reversal. The results were consistent with the predictions of LINOP, that is, faster acquisition when the preceding schedules were long, and when the following schedules were short. Grace concluded that a simple linear operator model can describe the

acquisition of preference, and suggested a single process might underlie choice in steady-state and in transition.

Overall, although fewer studies on acquisition of choice have been conducted with concurrent chains than concurrent schedules, there have been some notable results. Most importantly, evidence is mixed regarding the status of the 'value hypothesis': Some results favour the assumption that initial-link response allocation is determined by the relative value of the terminal links, with value a function of reinforcer delays (Grace, 2002 a, b; Mazur et al., 2001), whereas others pose a challenge to value as an explanatory construct (Grace & Nevin, 1999). However, the success of the LINOP model in describing results from Grace's (2002a) study in quantitative terms, together with Mazur's (1992) model for concurrent schedules, underscores the utility of the linear-operator approach. Next we consider studies which have used a different experimental design to study the acquisition of choice, specifically one in which contingencies change unpredictably across sessions.

Pseudo-Random Binary Sequence

Hunter and Davison (1985) used a concurrent variable-interval schedule where programmed reinforcer ratios changed from session to session according to a pseudo-random binary sequence (PRBS). As the name suggests, a PRBS consists of a pair of values that occur in an essentially random order such that there is zero correlation across presentations. As applied to concurrent schedules, the PRBS indicates which alternative (left or right) has the richer schedule in any given session (e.g., 4:1 or 1:4 reinforcer ratio). After pigeons had been exposed to this procedure, Hunter and Davison found that response ratios in the current session were a function of

reinforcers in the current session and previous sessions, but their analysis of a linear transfer function suggested that the effect of abrupt changes in the reinforcer ratio would be complete after approximately 5 sessions. This is substantially shorter than steady-state procedures, where training typically lasts for 15 to 30 sessions and training is usually terminated by response rates meeting a predetermined criterion. Thus, this procedure appears to offer a quick method for obtaining stable reinforcer sensitivity estimates and indicates that current response allocation is a function of reinforcement from different epochs.

Schofield and Davison (1997) examined the effects of previous sessions' reinforcer ratios on current performance using a concurrent-schedules design in which pigeons were exposed to three consecutive 31-step PRBS. In two experiments, using both non-independent and dependent scheduling, Schofield and Davison found that response allocation was largely determined by the reinforcer ratio in the current session. In addition, there was no effect of the size of the reinforcer ratios on the overall sensitivity to reinforcement or the sensitivity to reinforcement of the current session. Moreover, there was no evidence of a difference between the types of scheduling. Because the cumulative sensitivity to reinforcer ratios in the current and prior sessions was similar to that obtained in steady-state designs (approximately 0.80), Schofield and Davison suggested that the PRBS might offer a more efficient method of determining sensitivity to reinforcement.

Grace, Bragason, and McLean (2003) used a variation of the PRBS to examine concurrent-chains response allocation. In Experiment 1 one terminal-link schedule was FI 8 s while the other varied between FI 4 s and FI 16 s according to a PRBS. Grace et al. found response allocation to be most sensitive to the current session with some effect from previous sessions. This was similar to research using concurrent-

schedules (Hunter & Davison, 1985). However, reinforcer sensitivity values accumulated across sessions were lower than comparable steady-state research (Grace, 1994). In their Experiment 2, one terminal link was again fixed at FI 8 s, while a unique value was used for the other schedule, sampled from between FI 2 s and FI 32 s. The sensitivity values were similar to Experiment 1, but generalized matching plots suggested a clustering of data points into groups. For all pigeons, sensitivity values were greater when the changing alternative was greater than 8s compared to when it was less than 8s. Grace et al. suggested this indicated a categorisation process where pigeons were possibly discriminating which terminal link provided the shorter delay, rather than choice being determined by the ratio of the delays.

Decision Model

Grace and McLean (2006) assessed response allocation when the degree of variation across sessions was manipulated in a concurrent-chains design similar to Grace, Bragason, and McLean (2003). Specifically, they compared the sensitivities in a "minimal variation" condition, where one terminal link was constant and the other changed according to a PRBS, with a "maximal variation" condition that had unique terminal-link schedule pairs in each session. In both conditions, the average log immediacy ratio for sessions in which the richer terminal link was associated with the left (or right) key was log (2) (or log [1/2]). Each condition consisted of three PRBS presentations (93 sessions), and the order was counterbalanced. They found that response allocation tracked the current-session immediacy ratio in both conditions,

and across pigeons, with no systematic difference in sensitivity. Moreover, in the maximal-variation condition, generalized-matching plots (i.e. log response ratio vs. log immediacy ratio) appeared sigmoidal for one pigeon, suggesting a categorical discrimination process. However, for other pigeons, response allocations were approximately linear, consistent with the generalized-matching law.

Grace and McLean (2006) proposed a 'decision model' that could account for both linear and non-linear response allocation (i.e., generalized matching and categorical discrimination). The Decision Model assumes that response allocation is determined by the relative response strength of the initial-link schedules (that is, the relative propensity to respond to each alternative). Response strength for a particular initial link is updated after reinforcement has been obtained in the preceding terminal link, depending on the duration of the terminal-link delay relative to past delays to reinforcement. According to the Decision Model, subjects make a "decision" as to whether the preceding terminal-link delay was short or long. If the delay is judged as short relative to the history of reinforcement delays across both alternatives, response strength for the associated initial link increases, while if the delay is judged as long then response strength decreases. Changes in response strength increment and decrement according to a linear-operator rule (with parallel equations for left and right alternatives):

$$RS_{N+1} = RS_N + P_{'S'}(Max_{RS}\text{-}RS_N)\Delta - (1\text{-}P_{'S'})(RS_N - Min_{RS})\Delta \quad (21)$$

where, according to Equation 21, RS_{N+1} (predicted response strength for trial N+1) is determined by response strength on the previous trial (RS_N), modified by an additive (or subtractive) term, depending on whether the delay was judged as short (or long). Specifically, if the previous delay was judged as short (with probability $P_{'S'}$), the response strength increases by a constant fraction (determined by a learning rate parameter, Δ) of the difference between the maximum response strength (Max_{RS}) and current response strength. Conversely, if the previous delay was judged as long (with probability $1\text{-}P_{'S'}$), response strength decreases by a constant fraction of the difference between the current and minimum response strength (Min_{RS}). Whether a delay is classified as short or long depends on a comparison with the distribution of delays experienced across both alternatives. The model assumes a lognormal distribution of experienced delays with a mean equal to the log geometric mean. This allows a judgement criterion, the mean of the theoretical log normal distribution, to be calculated. The probability that a delay is judged "short" is the likelihood that a delay randomly sampled from the distribution is greater than the delay experienced in the preceding terminal link. The standard deviation (σ) is a parameter in the model and determines the accuracy with which delays are classified as short or long. Note that the delay that pigeons are equally likely to classify as short or long – the log geometric mean – corresponds to the bisection point that is typically obtained in temporal discrimination tasks (Church & Deluty, 1977; Stubbs, 1968).

Grace and McLean (2006) showed that the Decision Model could predict response allocation that conformed to generalized matching or categorical discrimination, depending on the value of σ. When σ was relatively low, classification decisions were accurate and response allocation was a nonlinear (sigmoidal) function of the log immediacy ratio. When σ was relatively large, decisions were less accurate,

and response allocation was approximately a linear function of the log immediacy ratio. They also showed that the Decision Model provided a reasonably good fit to the data for all their pigeons. The ability of the Decision Model to predict both response allocation patterns suggests that these two mechanisms could be fundamental processes involved in choice acquisition. Therefore, if the Decision Model can be generalized to steady-state and describe matching behaviour then it would be a legitimate contender with the existing models. Moreover, as the Decision Model assumes that the propensity to respond to the initial links, rather than the conditioned reinforcement of the terminal-link stimuli, is the mechanism for determining choice behaviour, this result would challenge the traditional assumptions of conditioned reinforcement being the molecular cause of choice behaviour.

Literature Review - Conclusion

Although existing steady-state models of choice can describe asymptotic behaviour, the question of what molecular processes underlie matching behaviour is unresolved. Acquisition research has traditionally used linear-operator models (Bush & Mosteller, 1951; Grace, 2002a; Lea & Dow, 1984; Mazur, 1992, 1997; Rescorla & Wagner, 1972). Such models generally calculate updated response strength as a constant proportion of the difference between current response strength and an asymptote. The linear-operator process has been successful at describing simple acquisition and also choice in transition (Grace, 2002a). Moreover, models using similar methods, moving averages based on current and past reinforcement, and feedback functions that inform the changing value of response strength, have also had some success (Myerson & Hale, 1988; Myerson & Miezin, 1980; Staddon, 1977).

Some researchers have suggested that models need to incorporate cognitive processes, such as memory and decision mechanisms, in order to provide a complete account of operant choice (Gallistel & Gibbon, 2000). Thus, the Decision Model's inclusion of both a linear-operator updating mechanism and a memory decision process represents a compromise. However, as a comprehensive description for choice the Decision Model is incomplete. It is unable to account for temporal context effects on choice, which are predicted by steady-state models such as DRT (Fantino, 1969). Moreover, the Decision Model describes response allocation within an individual session but not changes in response strength across sessions, which is incompatible with steady-state designs where response strength is assumed to gradually build and stabilise after several sessions. Therefore, although there is evidence to suggest the Decision Model of choice can describe acquisition in transition, it is only a partial model of choice.

The Present Research

The present research is an attempt to extend the Grace and McLean (2006) Decision model to a viable account of steady-state choice behaviour. As prior sections have shown, steady-state models of choice have been traditionally challenged to describe the effects of changes in initial-link duration on response allocation (Fantino, 1969, Grace, 1994, Mazur, 2000). Because it contains no way for initial-link duration to affect choice, the original Grace and McLean Decision model (DM) is incomplete as a model for concurrent chains.

This thesis presents an extension to the DM that attempts to describe the effects of temporal context and transitions across sessions, and also derives a steady-state version of the extended model. The extension is presented as a series of

mathematical calculations and attendant assumptions. Together, these extensions are referred to as the extended decision model (EDM). The first extension proposes to address the issue of the effect of temporal context in concurrent chains. Specifically, this extension assumes the initial-link durations – the intervals between the onset of the initial links and terminal-link entry – also affect the decisions about whether a terminal-link delay is judged short or long. Reinforcement history is assumed to include the delays between all stimuli correlated with food, including the onset of the initial links and terminal-link entry, as well as between terminal-link onset and food delivery. Although these delays are experienced as quite different events this mechanism allows for the subject's experience of the total time to reinforcement to be included in the EDM calculation of response allocation. Moreover, the inclusion of both initial - and terminal-link delays in the Decision Model calculation is similar to existing models of choice, such as DRT (Fantino, 1969; Squires & Fantino, 1971) and HVA (Mazur, 2001), that assume total time and time from stimuli onset to reinforcement, respectively, influence response allocation. By including the initial-link duration, the EDM has a potential mechanism for describing temporal context effects.

The original DM assumed the criterion to be the mean of the theoretical log normal distribution of experienced delays. This is used to compare the just experienced delay so a judgement can be made whether this was 'short' or 'long'. This process can be formally expressed as the probability of short categorisations p_s:

$$P_s = 1 - \Phi(\text{Log D}, Log\ C, \sigma) \qquad (22)$$

where Φ is the cumulative normal distribution, *Log C* (criterion) the log normal mean, and σ the standard deviation, evaluated for the preceding delay including initial and terminal-links (Log D). This calculation computes P_s as the area under the distribution of experienced delays to the right of the log of the preceding delay. This value is used to calculate the change in response strength using the linear-operator component of the DM.

The inclusion of both links has the effect of influencing the criterion (that is, the mean of the reinforcement history distribution) so that it varies directly with both initial-link and terminal-link duration. This creates a unique prediction – as initial-link duration increases and terminal link duration remains constant, response allocation follows a bitonic rather than monotonic function. This is in conflict with existing research that has examined varying the duration of the initial link, specifically, initial-link effect; increasing initial-link duration when terminal-link duration is constant linearly decreases response allocation (Fantino, 1969). Figure 1 shows representative predictions of the extended decision model for preference between FI 10 s and FI 20 s terminal links, as the initial-link duration increases from 5 s to 30 s. Over most of the range of initial-link durations – between 10 s and 30 s – the predicted preference becomes less extreme as the initial links increase.

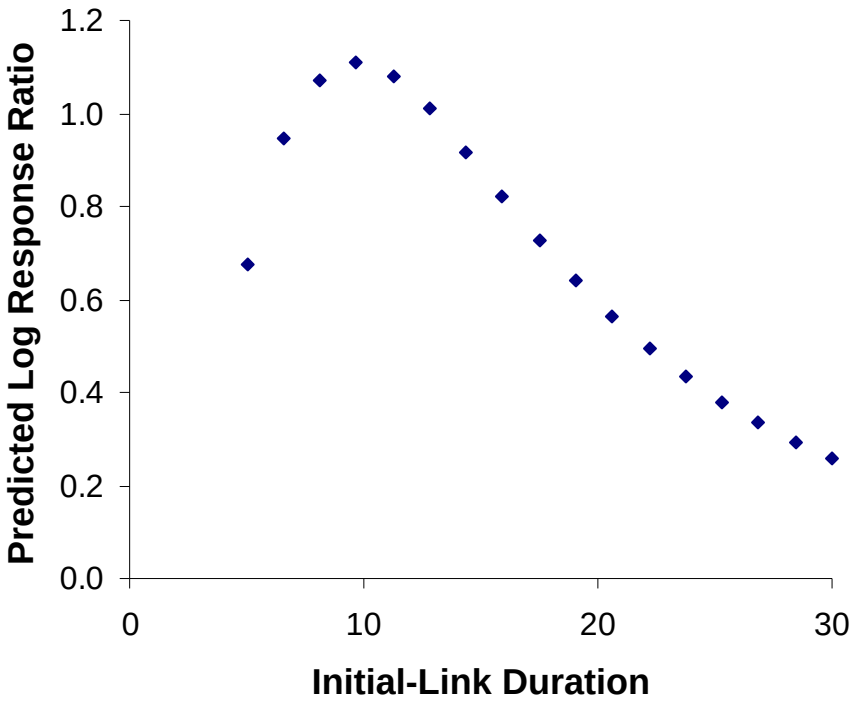


Figure 1. Preference for the shorter (FI 10 s) terminal link predicted by the extended version of Grace and McLean's (2006) Decision Model, as the initial-link duration increases from 5 s to 30 s. For the predictions in Figure 1, parameter values for the decision model were $\sigma = 0.15$, $Max_{RS} = 1.00$, $Min_{RS} = 0.01$.

The reason is that preference becomes less extreme is that the relative probability of a short decision for the terminal links, that is, P('short' | 10 s) / P('short' | 20 s), decreases as the initial-link duration increases. As a result, including the time spent in the initial links affects the calculation of the criterion and allows the extended decision model to account for response allocation changes due to varying initial-link duration. However, Figure 1 also shows that for short initial-link durations, there is a range over which preference become less extreme as the initial links decrease. Thus, the overall function relating preference to initial-link duration is

bitonic. This occurs because the effect of the shorter terminal-link delay on value changes in relation with the criterion. The $P_{'S'}$ is categorized 'short' declines as the initial-link duration becomes smaller relative to the terminal-link midpoint. At approximately the terminal-link midpoint, the discrimination of the shorter terminal-link begins to decline with increases in initial-link duration, creating a bitonic effect. Figure 2 illustrates the bitonic effect as a combination of the changing $P_{'S'}$ for the two alternatives.

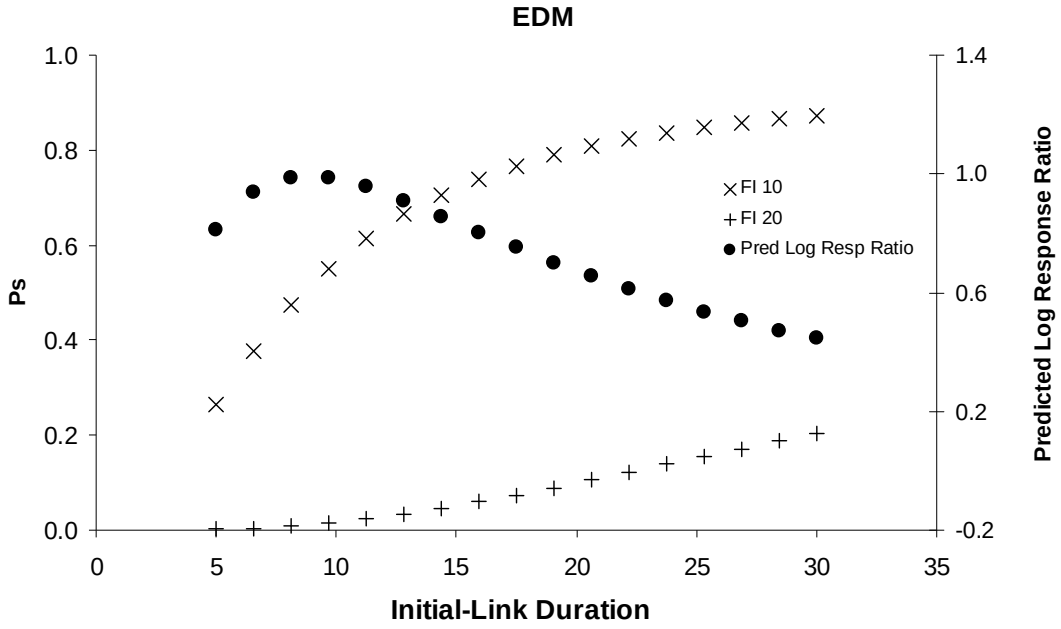


Figure 2. Log initial-link response allocation (filled circles, right axis) predicted by the EDM as a function of initial-link duration. Data points indicated by x's and +'s show the probabilities of a terminal-link delay judged short ($P_{'s'}$, left axis) relative to the criterion for the FI 10 and FI 20 schedules, respectively.

This bitonic function appears robust with respect to variation in the parameter values, and represents a novel prediction of the Extended Decision Model (EDM)

compared to existing models such as CCM (Grace, 1994), HVA (Mazur, 2001), and DRT (Fantino, 1969) that predict choice is a monotonic decreasing function of initial-link duration.

Grace and McLean (2006) updated their model after each experienced terminal-link delay. This approach is appropriate when all experiences are equally weighted, however research suggests that response allocation is a function of previous and current reinforcement histories, where more recent experiences are given greater weight than temporally distant experiences (Mazur, 1987). Therefore, in an attempt to describe the effect of temporally-distant experiences, an approach is needed to describe the changing influence of previous delays on the criterion. Research attempting to model the experience of changing delays on response allocation has used temporal weighting rules (Mazur, 1996), while other researchers have assumed an exponentially-weighted moving average (Killeen, 1981, 1994). This is the approach taken here: the criterion is assumed to change based on the experience of successive presentations of delays between reward-correlated stimulus transitions according to an exponentially-weighted moving average:

$$\log C_{N+1} = \beta\left(\log D_N\right) + \left(1-\beta\right)\log C_N \tag{23}$$

where $\log C_N$ and $\log C_{N+1}$ are the criterion values after stimulus transitions N and N+1, respectively, $\log D_N$ is the Nth stimulus-transition delay, and β is a parameter that determines how much weight to give to the most recent delay. Note that N does not correspond to cycle number, because the criterion is updated twice per cycle – first after terminal-link entry (i.e., the initial-link -> terminal-link onset delay), and

then again after food delivery (i.e., the terminal link -> food delay). With the addition of Equation 23, the EDM can be applied to situations in which the criterion might shift within sessions, for example, in which the initial- or terminal-link schedules are changed during a session so that earlier initial- and terminal-link delays recede in importance whereas new delays gain influence.

The original DM assumes as response allocation is "reset" at the beginning of each session. This assumption may be appropriate for PRBS designs, but is incompatible with steady-state research where changes in response allocation cumulate across sessions, gradually approaching asymptotic levels. Thus, to describe the acquisition across process across sessions some degree of 'carry-over' is assumed. Specifically, response strength at the beginning of a session is assumed to be a constant proportion of the change in response strength from the previous session added to response strength of the previous session:

$$RS_{startN+1} = RS_{startN} + (RS_{endN} - RS_{startN})\Delta s \qquad (24)$$

where RS_{start} and RS_{end} are response strength predictions at the start and end of the session (subscripted N or N+1) respectively, and Δs is a learning rate parameter. Thus, the EDM is an event-based rather than time-based model. With the addition of Equation 24, the EDM can describe both within- and between-session learning. Note that Δs is assumed to be generally less than 1, so that response strength at the start of session N+1 will have regressed back towards the response strength at the start of the previous N. The assumption in the EDM that a fraction of the change in response strength during a session carries over to the next session provides a natural description for spontaneous recovery in choice behaviour. For example, Mazur (1995, 1996)

found that when the proportion of reinforcers delivered by an alternative was changed midway through a session (e.g., from 10% to 90%), pigeons' response allocation would shift (e.g., from 10% to 75%), but at the start of the next session would have reverted to an earlier percentage (e.g., 45%). Mazur proposed that this effect, which resembles spontaneous recovery, was described by assuming that the response strengths at the start of a session were determined by a weighted average of the several previous sessions. The EDM can predict the same result through a different but arguably simpler mechanism.

We have now arrived at a position to derive a steady-state version of the DM. First, it is assumed that response strengths reach their asymptotic values, which is described as the weighted average of the maximum and minimum response strengths, with weights determined by the probability of a short decision:

$$\mathrm{RS}_{\mathrm{asymp}} = \mathrm{P_S}\mathit{Max}_{\mathrm{RS}} + (1 - \mathrm{P_S})\mathit{Min}_{\mathrm{RS}} \qquad (25)$$

where $\mathrm{RS}_{\mathrm{asymp}}$ is asymptotic response strength, $\mathrm{P_S}$ is the probability of choosing short, $\mathit{Max}_{\mathrm{RS}}$ is the maximum response strength, and $\mathit{Min}_{\mathrm{RS}}$ is the minimum response strength. Equation 25 shows that the predicted asymptotic response strength for each alternative is a combination of the weighted average of $\mathit{Max}_{\mathrm{RS}}$ and $\mathit{Min}_{\mathrm{RS}}$, with the weights given by the probabilities that the delay is judged short or long, respectively. Therefore, applying the above calculation to each alternative, the predicted response ratio can be calculated:

$$\log\left(\frac{B_L}{B_R}\right) = \frac{RS_{asympL}}{RS_{asymp\,R}} = \frac{P_{sL}Max_{RS} + (1 - P_{sL})Min_{RS}}{P_{sR}Max_{RS} + (1 - P_{sR})Min_{RS}} + \log b \quad (26)$$

where B is initial-link response rate and subscripts L, R indicate the choice alternatives, log *b* is a bias parameter and the remaining parameters have been described in previous paragraphs. Thus, the Equation 26 assumes steady-state behaviour to be based on the relative asymptotic response rates, which are the result of the combined weighted average of minimum and maximum response rates for each alternative. Calculating separate response strength for each alternative allows the separate parameters from each alternative to affect steady-state response prediction. Moreover, creating a steady-state version of the decision model allows comparisons with other models of steady-state behaviour. Furthermore, as the EDM assumes a novel process for determining value, the propensity to respond rather than conditioned reinforcement provides another point of difference between the EDM with other concurrent-chain models.

Purpose

The purpose of the present research is an examination of the proposed EDM in experiments that have traditionally challenged steady-state models of choice – temporal context. Specifically, this thesis investigated the effect of changing the initial- and terminal-link durations in three types of procedure; quasi-random, short or long initial- or terminal-link durations, and when immediacies change each session

but follow a systematic ascending and descending pattern. These investigations are separated into chapters that examine changing either the initial- or terminal-link durations: Chapter 2 contains Experiments 1-3 that assess the effect of changing initial-link duration, while Chapter 3 contains Experiments 4 and 5 that examine the effect of changing the terminal-link duration. In Chapter 4, the steady-state version of the EDM is compared against existing steady-state models of choice when fitted to over a dozen published archival data sets (Grace, 1994). The general discussion in Chapter 5 will draw conclusions from these results and suggest some possible implications.

Chapter 2: Initial Link Experiments

Experiment 1

Introduction

Previous concurrent-chains studies have used steady-state procedures where training with a given pair of terminal-link schedules continues until response allocation in the initial links has stabilised. Typically this requires 20 or more sessions, after which the terminal-link schedules are changed and training begins in a new condition (for example, Fantino & Davison, 1983; Grace, 2002a). However, studies have shown response allocation can track changes in terminal-link schedules that occur unpredictably across sessions (Grace et al., 2003; Grace & McLean, 2006). For example, Kyonka and Grace (2007) exposed pigeons to a concurrent-chains procedure where the terminal-link schedules in each session were either fixed-interval (FI) 10 s FI 20 s or FI 20 s FI 10, determined by a pseudo-random binary series, and the initial link was a variable-interval (VI) 8-s schedule that arranged equal access to the terminal links. After pigeons had received about 50 sessions of training, response allocation stabilised about midway through each session, showing strong sensitivity to the terminal-link delays in the current session with virtually no influence of prior sessions.

This experiment examines response allocation behaviour in a rapid-acquisition concurrent-chains procedure with a constant pair of terminal-link schedules (FI 10 s and FI 20 s). The rapid-acquisition procedure was used in this experiment to examine

response allocation in choice in transition. Training began with a VI 10s initial-link schedules and a constant pair of Fixed Interval 8 s terminal-link schedules. When subjects began showing sensitivity to the immediacy ratio in the current session the initial-link schedule changed, following a systematic progression between 0.01 s and 30 s, in an ascending and descending series. This experiment was designed to investigate whether the stable level of response allocation reached within sessions would vary with initial-link duration, and whether the data would be a decreasing monotonic function (as predicted by current models for steady-state choice), or a bitonic function as predicted by the extended decision model.

Method

Subjects

Six pigeons of mixed breed, numbered 181, 182, 183, 184, 185 and 186 served as subjects and were maintained at 85% of their free-feeding weight ±15 g through appropriate post-session feeding. Subjects were housed individually in a vivarium with a 12h:12h light/dark cycle (lights on at 0600), with water and grit freely available in the home cages. All pigeons were experienced with a variety of experimental procedures.

Apparatus

Four standard three-key operant chambers, 32 cm deep x 34 cm wide x 34 cm high, were used. The keys were 21 cm above the floor and arranged in a row. In each chamber there was a house light located above the centre key that provided general illumination, and a grain magazine with an aperture centred 6 cm above the floor.

The magazine was illuminated when wheat was made available. A force of approximately 0.15 N was necessary to operate each key. Each chamber was enclosed in a sound-attenuating box, and an attached fan provided ventilation and white noise. Experimental events were controlled and data recorded through a microcomputer and MEDPC® interface located in an adjacent room.

Procedure

All pigeons started training immediately in a concurrent-chains procedure. The house light provided general illumination at all times except during reinforcer delivery. With few exceptions, sessions were run daily and at approximately the same time (12:00h). Sessions ended after 72 initial- and terminal-link cycles or 70 min, whichever occurred first. At the start of a cycle, the side keys were illuminated white to signal the initial links. Terminal-link entry was assigned quasi-randomly to the left or right with the constraint that in every 6 cycles, 3 entries occurred to each terminal link. An initial-link response produced an entry into a terminal link provided that: (a) it was made to the pre-selected key; (b) an interval selected from the initial-link schedule had timed out; and (c) a 1-s changeover delay (COD) was satisfied— at least 1 s had elapsed following a response to the key for which terminal-link entry was arranged. The COD was in effect throughout a cycle, except during reinforcement, and was reset prior to each cycle, so that the first response to either key in each cycle was considered a changeover to that alternative.

A single VI schedule operated during the initial links. The initial-link VI schedule contained 12 intervals constructed from an exponential progression (Fleshler & Hoffman, 1962), except for the 0.01 interval which was a fixed 0.01s interval. Separate lists of intervals were maintained for cycles in which the left or right

terminal link had been selected, and were sampled without replacement so that all 12 intervals would be used three times for both the left and right terminal links each session. The 12 intervals for each delay were logarithmically spaced so that their arithmetic average was a linear function of the mean schedule duration. This created a wide distribution of intervals for each delay, typically in the order of three times the size of each initial-link schedule. For example, at 8 s the maximum interval was 27.8 s, providing ample time for subjects to respond on both keys. When a terminal link was entered, the colour of the side key was changed (left key to red, right key to green) while the other key was darkened. Terminal-link responses were reinforced according to FI schedules. The terminal-link schedules were always FI 10 s and FI 20 s, but the location of the richer schedule varied across sessions according to a 31-step quasi-random binary series similar to that used by Hunter and Davison (1985). When a response was reinforced, all lights in the chamber were extinguished, the grain magazine was raised and illuminated for 3 s, and then the next cycle began.

The experiment consisted of two phases, which varied only in terms of the initial-link schedule. In Phase 1, the initial-link schedule was always VI 10 s. This consisted of period of time where the subject was able to key peck at two lighted keys. These responses were recorded when the key was struck with enough force, and after a switch, the COD was satisfied. The first phase was terminated individually for each pigeon when regression analyses showed that response allocation during the last 20 sessions showed strong sensitivity to the immediacy ratio in the current session (that is., a in Equation 4 >1.5 and negligible position bias (log b in Equation 4 < 0.10). Training consisted of 150, 193, 92, 193, 78 and 87 sessions for pigeons 181 to 186, respectively. Initial-link responses rates were determined to stabilise by visual inspection of the daily choice allocations.

In Phase 2, the initial-link changed across sessions according to an ascending and descending series. The series varied from 0.01 to 30 s in 17 equally spaced steps. The schedule values were: 0.01, 1.88, 3.75, 5.63, 7.50, 9.38, 11.25, 13.13, 15.00, 16.90, 18.75, 20.63, 22.50, 24.38, 26.25, 28.13 and 30 s. As the initial-link schedule had been VI 10 s in Phase 1, training began in the middle of the descending (9.38, 7.5, 5.63, etc.) or ascending (11.25, 13.13, 15.00, etc.) series, counterbalanced across birds. When the limit of either series was reached, the direction was reversed and the other sequence began with the same values in opposite order (for example, 24.38, 26.25, 28.13, 30, 28.13, 26.25, 24.38, etc.). For sake of convenience, 0.01 s and 30 s were assigned to the descending and ascending series, respectively. Statistical analyses that compared the ascending and descending series were based only on those values that were common to both series (that is, 0.01 s and 30 s were excluded). Training in Phase 2 continued until all pigeons had completed the full ascending and descending series at least two times each. As some pigeons completed training in Phase 1 earlier than others, the number of sessions in Phase 2 varied. The total number of sessions completed in Phase 2 was 75, 79, 166, 80, 165 and 167 sessions for pigeons 181 to 186, respectively.

Results

To assess the relationship between response allocation and the immediacy ratios in the current and prior sessions, the generalized-matching model was used:

$$\log\frac{B_{0L}}{B_{0R}} = a_0\log\frac{1/D_{0L}}{1/D_{0R}} + a_1\log\frac{1/D_{1L}}{1/D_{1R}} + a_2\log\frac{1/D_{2L}}{1/D_{2R}} + \ldots + \log b \quad (27)$$

where B and *D* refer to initial-link response rate and terminal-link delay, respectively, subscripted for choice alternative (L and R) and lag (0 through 4; 0 = current session). The parameters a_0 ... a_4 quantify sensitivity to reinforcer immediacy (that is., reciprocal of delay) at each lag, and log *b* is a bias parameter.

Equation 27 was applied to the data for individual subjects from the last 30 sessions of Phase 1, and all of the sessions from Phase 2. The upper panel of Figure 3 shows results when responses during the second half of each session were analysed with Equation 27. For all subjects in both phases, Lag 0 coefficients were positive and statistically significant, and none of the higher lag coefficients reached significance using a multiple-regression analysis. The lags were calculated from the log immediacies from corresponding sessions. For example, lag 0 came from the terminal-link immediacy from the current session (D_0), lag 1 came from terminal-link immediacy from the most recent session (D_1), lag 2 came from the session 2 sessions prior to the current session (D_2), etc . The following figure shows the sensitivity coefficients across subjects for each lag.

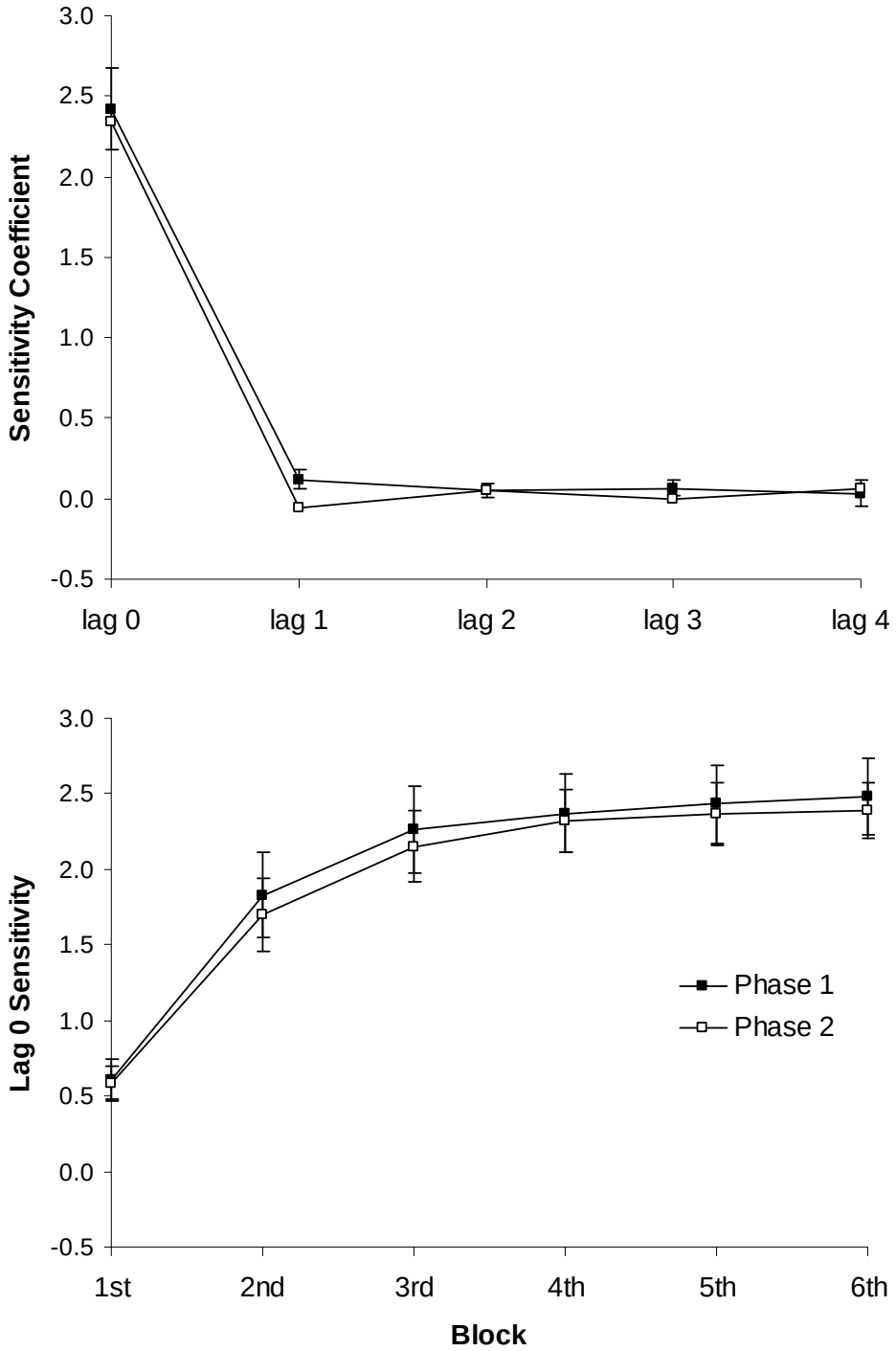
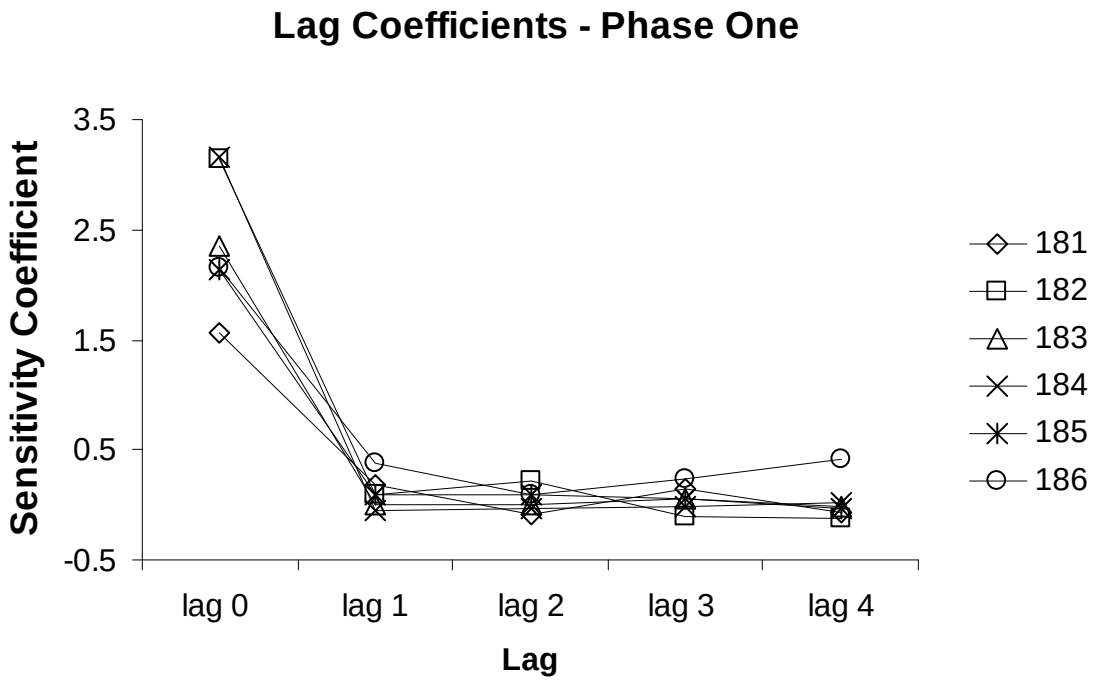


Figure 3. The upper panel shows average sensitivity coefficients from the last half of each session for Lag 0 through to Lag 4 log immediacy ratios, for both Phase 1 (filled symbols) and Phase 2 (unfilled symbols). The lower panel shows average Lag 0 sensitivity coefficients determined separately for each block of 12 trials within sessions, for both Phase 1 (filled symbols) and Phase 2 (unfilled symbols). The dashed line represents zero sensitivity. Bars indicate +/- 1 *SE*.

A repeated-measures analysis of variance (ANOVA) found a significant effect of lag, $F(4,20) = 110.88$, $p < .001$, but the effects of phase and the lag x phase interaction did not reach significance. Averaged across subjects, Lag 0 sensitivity to immediacy was 2.42 ($SE = 0.26$) in Phase 1 and 2.35 ($SE = 0.20$) in Phase 2 for the second half of the session. Figure 3 shows that the responding in the second half of the sessions was determined by the immediacy ratio in the current session, with little or no effect of immediacy ratios from previous sessions. To examine the acquisition of preference within sessions, Equation 25 was applied to data from each of the six blocks (session 6ths, or 12 cycles per block). The lower panel of Figure 3 shows the results averaged across subjects. A repeated-measures ANOVA found a significant effect of block, $F(5,25) = 106.14$, $p < .001$, but the effects of phase and the block x phase interaction did not reach significance.

Figure 3 show that response allocation reached stability approximately halfway through the session, consistent with previous studies using similar rapid-acquisition procedures (Grace et al., 2003; Kyonka & Grace, 2007, 2008). Averaged across subjects, the Lag 0 sensitivity to immediacy in the second half of the session was 2.43 [$SE = 0.26$] in Phase 1 and 2.36 [$SE = 0.20$] in Phase 2. Figures 4 and 5 show the individual Lag coefficients. The individual results mirror the group results suggesting this was a consistent phenomena across subjects.



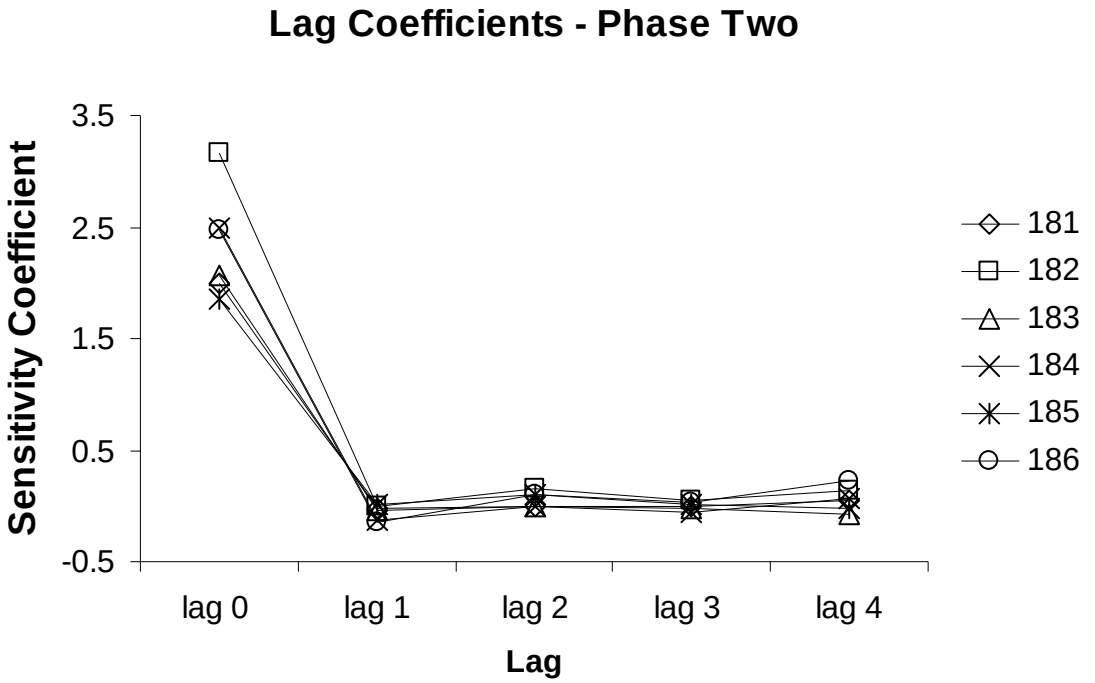


Figure 4. The upper panel shows the average individual sensitivity coefficients for Lag 0 through to Lag 4 log immediacy ratios in the second half of the session in Phase 1, and the lower panel shows the average individual sensitivity coefficients for lag 0 through to lag 4 immediacy ratios for Phase 2.

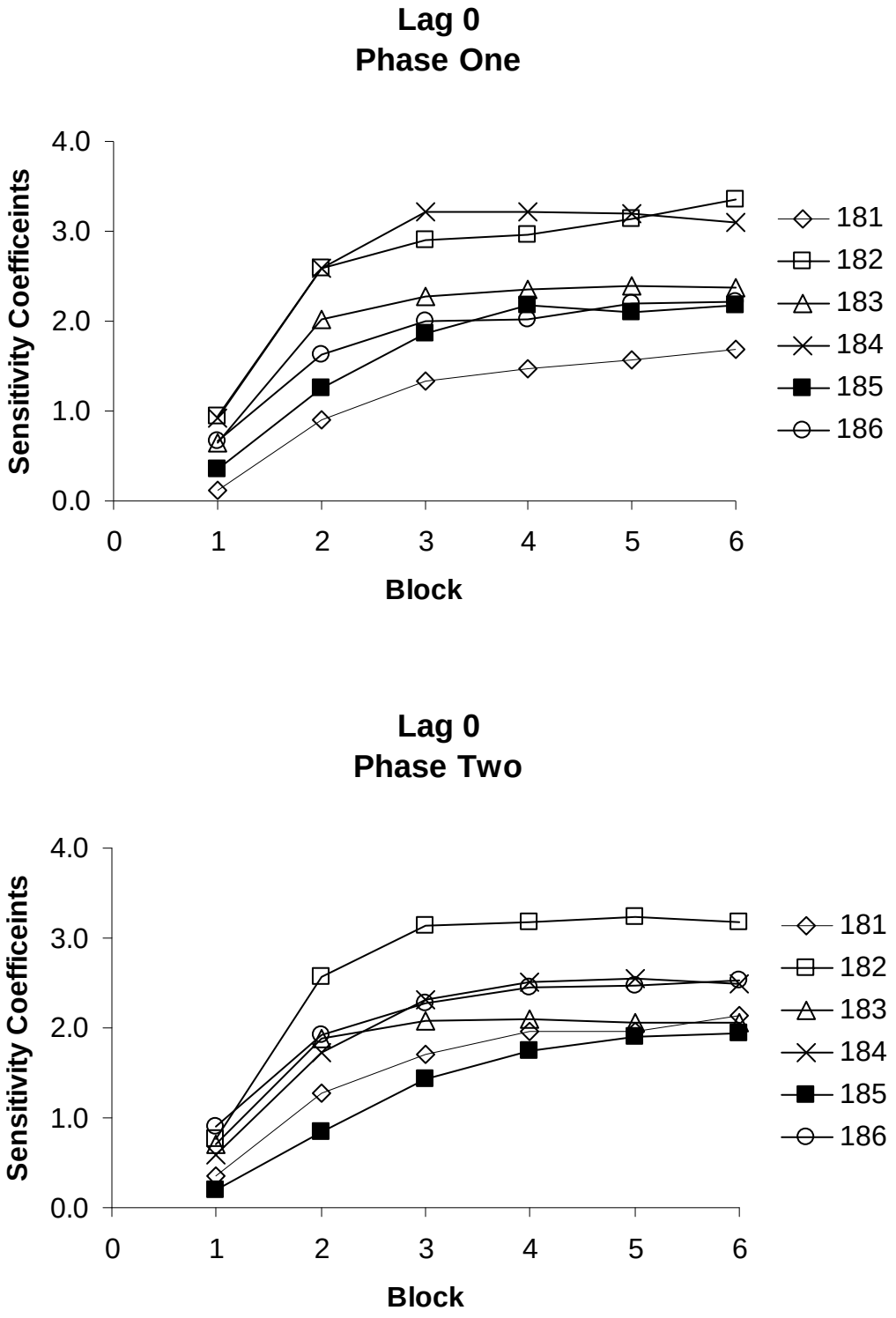


Figure 5. The upper panel shows average individual Lag 0 sensitivity coefficients determined separately for each block of 12 trials for phase one, and the lower panel shows average individual Lag 0 sensitivity coefficients for each block of 12 trials for phase two.

The individual results show during the last 30 sessions of Phase 1 and throughout Phase 2 individual response allocation showed strong sensitivity to the immediacy ratio in the current session, consistent with prior studies. This suggests sensitivity to initial-link duration appeared to stabilise within session despite initial-links changing each session.

Figure 6 shows the obtained initial-link duration and response allocation plotted as a function of Phase two programmed initial-link duration. These plots show the correspondence between programmed and obtained initial-link duration and the bitonic effect predicted by the EDM (see Figure 2).

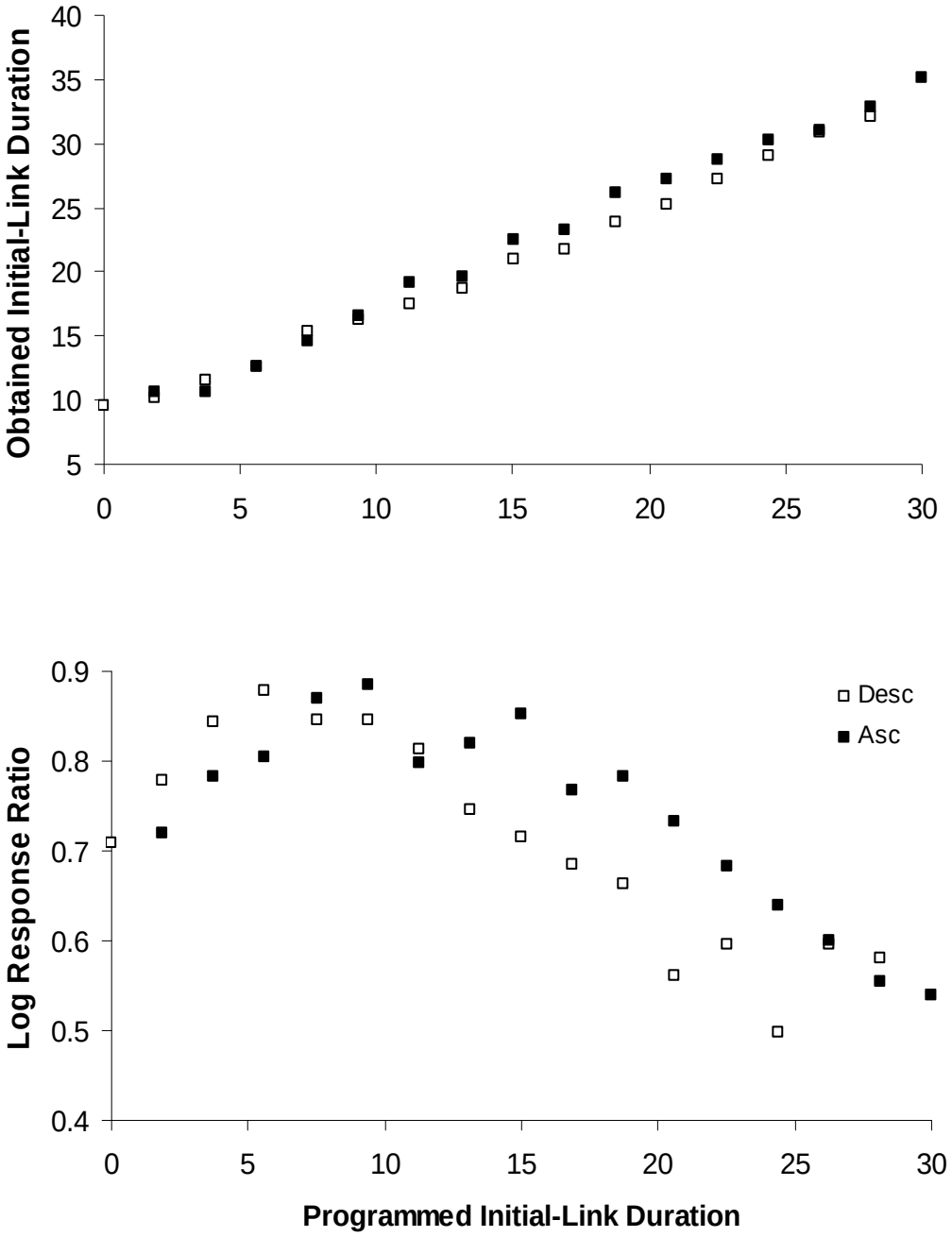


Figure 6. The upper panel shows average obtained initial-link duration plotted as a function of the programmed value, for both ascending (filled symbols) and descending (unfilled symbols) series in Phase 2. The bottom panel shows average obtained initial-link response allocation for the FI 10 s terminal links from the second half of the session, plotted as a function of programmed initial link duration for both ascending (filled symbols) and descending (unfilled symbols) series.

The upper panel of Figure 6 shows the average obtained average time spent in the initial links per cycle as a function of programmed initial-link duration in Phase 2 across subjects. Obtained initial-link duration increased linearly with programmed duration: the equation $y = 0.87x + 8.36$ accounted for over 99% of the variance across series. A repeated-measures ANOVA confirmed the effect of programmed initial-link value, $F(14,70) = 410.82$, $p < .001$, although obtained initial link durations were greater for the ascending than the descending series, $F(1,5) = 7.33$, $p < .05$. The interaction was not significant. These results show that obtained initial-link duration increased linearly with programmed duration, and that there was a minimum averaged obtained duration of approximately 8.36 s. This minimum duration occurred because of the dependent scheduling arrangement used in the initial links, which required the pigeon to make responses to the non-preferred alternative.

To examine whether response allocation depended on initial-link duration, the data from the second half of each session were analysed, where response allocation had stabilised. Bias estimates (log b in Equation 27) were subtracted from log initial-link response ratios so sessions with FI 10 s FI 20 s and FI 20 s FI 10 s could be compared on a common scale. Using the absolute log initial-link response ratio, replications of each programmed initial-link schedule value was averaged across subjects for each series.

The bottom panel of Figure 6 shows the average log response allocation as a function of programmed initial-link duration, for both ascending and descending series in Phase 2. Response allocation was a bitonic function of initial-link duration for both series. A repeated-measures ANOVA found a significant effect of initial-link duration, $F(14,70) = 16.94$, $p < .001$ and a significant interaction between initial-link

duration and series, $F(14,70) = 3.49$, $p < .001$. The main effect of series was not significant. Planned polynomial contrasts found significant linear, quadratic, and cubic trends for initial-link duration, $F(1,5) = 133.61$, $p < .001$, $F(1,5) = 7.90$, $p < .05$, and $F(1,5) = 36.60$, $p < .01$, respectively. The linear trend represents the classic "initial-link effect": overall response allocation became less extreme as initial-link duration increased. The quadratic and cubic components confirm that the non-monotonicity of the function in the bottom panel of Figure 6 was significant. As programmed initial-link duration increased from 0.01 s, response allocation became more extreme at first, but then decreased as initial-link duration increased. The cubic component resulted from the minimum preference during the descending series occurring at 24.38 s rather than at the end of the series.

The significant interaction in the group data between initial-link duration and sequence occurred because for relatively long initial-link durations, response allocations during the ascending series tended to be greater than during the descending series, whereas the opposite was obtained for relatively short durations. In effect, the bitonic function for the ascending series as a whole was shifted to the right compared with the descending sequence. This horizontal displacement suggests that a possible hysteresis or lag effect was present based on which series had previous sessions that were more extreme: For relatively long initial-link durations, the more extreme preferences for the ascending compared to descending series reflected the possible influence of initial-link durations from the preceding sessions, which were relatively shorter for the ascending series. Conversely, for relatively short initial-link durations, the less extreme preference evident in the ascending series could have occurred because the initial-link durations from preceding sessions were shorter. Thus, although there was no effect of the immediacy ratio from the prior sessions on

response allocation, there was an effect of initial-link duration. Individual obtained initial-link response allocation was plotted as a function of programmed initial link duration to investigate whether the bitonic effect was evident by visual inspection in the data for individual subjects.

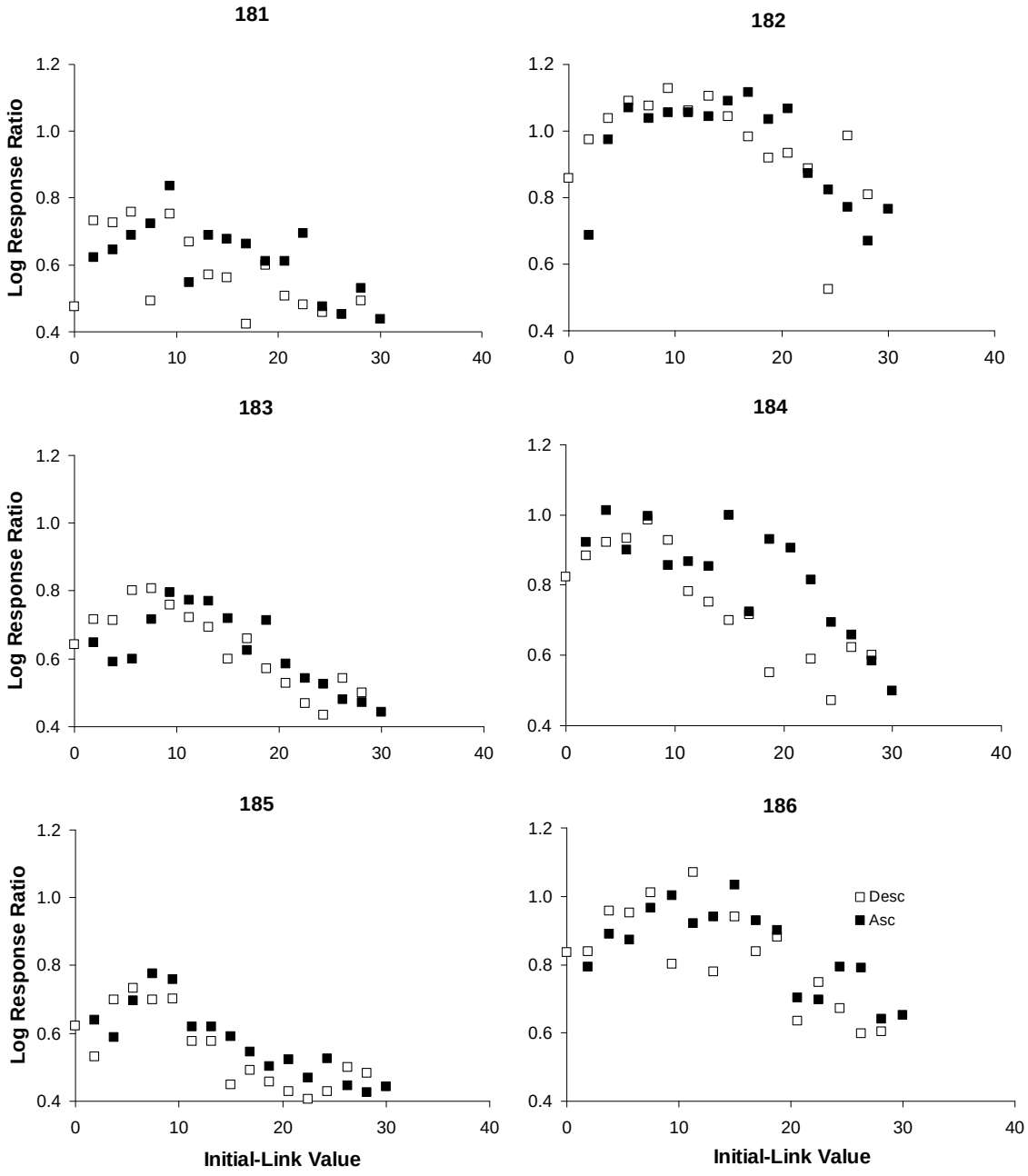


Figure 7. These plots show the individual obtained initial-link response allocation for the FI 10 terminal-links from the second half of the session, plotted as a function of programmed initial-link duration for both the ascending (filled black squares) and descending (unfilled black squares) series in Phase 2.

The individual plots of the response allocation across replications appear to show a possible bitonic effect for both the ascending and descending series. Across subjects, and for most series, the bitonic pattern is maintained: response allocations at the shorter initial-link durations are less that the individual maximums, whilst the largest delays have the lowest log response allocation ratios. However, compared with the group results, the bitonic pattern was more variable in the individual data. For example, subject 181 appears to have a function that is close to linear, whilst subjects 183 and 184 have bitonic patterns.

The EDM was then fitted to the individual subject data for Phase 1. The general process for EDM calculations in this thesis was as follows. The last 40 sessions from Phase 1 provided the raw data (note that this larger number of sessions was chosen because of the relatively large number of training sessions used in this experiment). The log ratio (left over right responses) was calculated for each session 12^{th} (6 cycles per session 12^{th}). This provided the 'obtained' data points (480 in this experiment) for each subject. Probabilities of a short categorisation ($P_{'s'}$, equation 20) were calculated for each alternative at each session 12^{th}, whilst the criterion for each session 12^{th} was calculated using the combination of the log initial- and terminal-link delays from both alternatives (equation 21). The log ratio of the DM calculations (equation 19) for each alternative was used as the 'predicted' session 12^{th} scores. The following DM parameters were allowed to vary freely (others are used in different experiments depending on the procedure): σ, log b, ΔL, ΔR, and β. The remaining DM parameters *Max* and *Min* were set at 1.0 and 0.01 for both alternatives, respectively. The EDM carry-over parameter Δs (equation 22) was set at 0.0 because Phase one used a quasi-random binary sequence of changing terminal-link values and subjects

received enough training to reduce carryover effects across sessions. The beginning DM value for both left and right alternatives were set at 0.5 because the model assumes a previous value before calculating the current value. The parameters ΔL and ΔR were chosen to vary because both terminal-link delays were changing each session. The parameters σ and β were allowed to vary because these describe the effects of changes in delay influencing the probability of a short categorisation, whilst log b was allowed to vary to fit any linear subject bias in the log response ratios. The values *Max* and *Min* were set at 1.0 and 0.01, yielding a possible range of predicted response allocation of four log units (100:1 to 1:100).

The percentage of variance in the obtained data accounted for by the predicted data was used as the target value that changes in the free parameters attempted to maximise. The maximisation process used the non-linear optimisation routine Solver found in the Excel® spreadsheet program. This routine maximized the target value within minimum convergence values of 0.0001 and a tolerance from the obtained values of 5%. A maximum of 1000 iterations was set to allow a sufficient number for Solver to calculate the highest target value. Additionally, initial starting parameter values were investigated to gain the highest target value score. Constrains for Δ and β were applied to so parameter values must be greater than 0.0001 and less than 2.0 to avoid possible inappropriate function or division by zero errors. These constraints were used to allow Solver to iterate solutions until the maximum fit between obtained and predicted was achieved. This produced the following parameter values.

Table 1

EDM VAC and Free Parameter Values for Individual Subjects Fitted to Session 12th Data in Phase 1.

	VAC	EDM Free Parameters				
	Session 12th	σ	Δ L	Δ R	Log b	β
181	0.60	0.19	1.00	0.00	0.28	0.88
182	0.69	0.16	1.11	0.27	0.08	1.26
183	0.90	0.18	1.00	0.30	0.08	1.38
184	0.90	0.15	0.90	0.32	0.00	1.10
185	0.83	0.19	0.33	0.22	0.03	0.06
186	0.56	0.14	0.05	0.50	-0.21	0.02
Average	0.75	0.17	0.73	0.27	0.04	0.78

Table 1 illustrates that the quality of EDM fits appeared to vary across subjects. For example, subjects 181, 182 and 186 VAC (average 0.62) was much lower than for subjects 183, 184 and 185 (average 0.88) Reasons for the individual differences in quality of fit are unclear. There appears no mediating EDM free parameter responsible for this variability suggesting the lack of fit is due to random variability.

Figure 8 shows the session 12th obtained log initial-link response ratios as a function of EDM predictions for the individual subjects in Phase 1. Slopes for the best fitting regressions are shown and the group average is close to 1.0 (mean = 1.05), suggesting that the EDM captured the major trends in the data. Averaged across subjects, EDM described 75% of the variance.

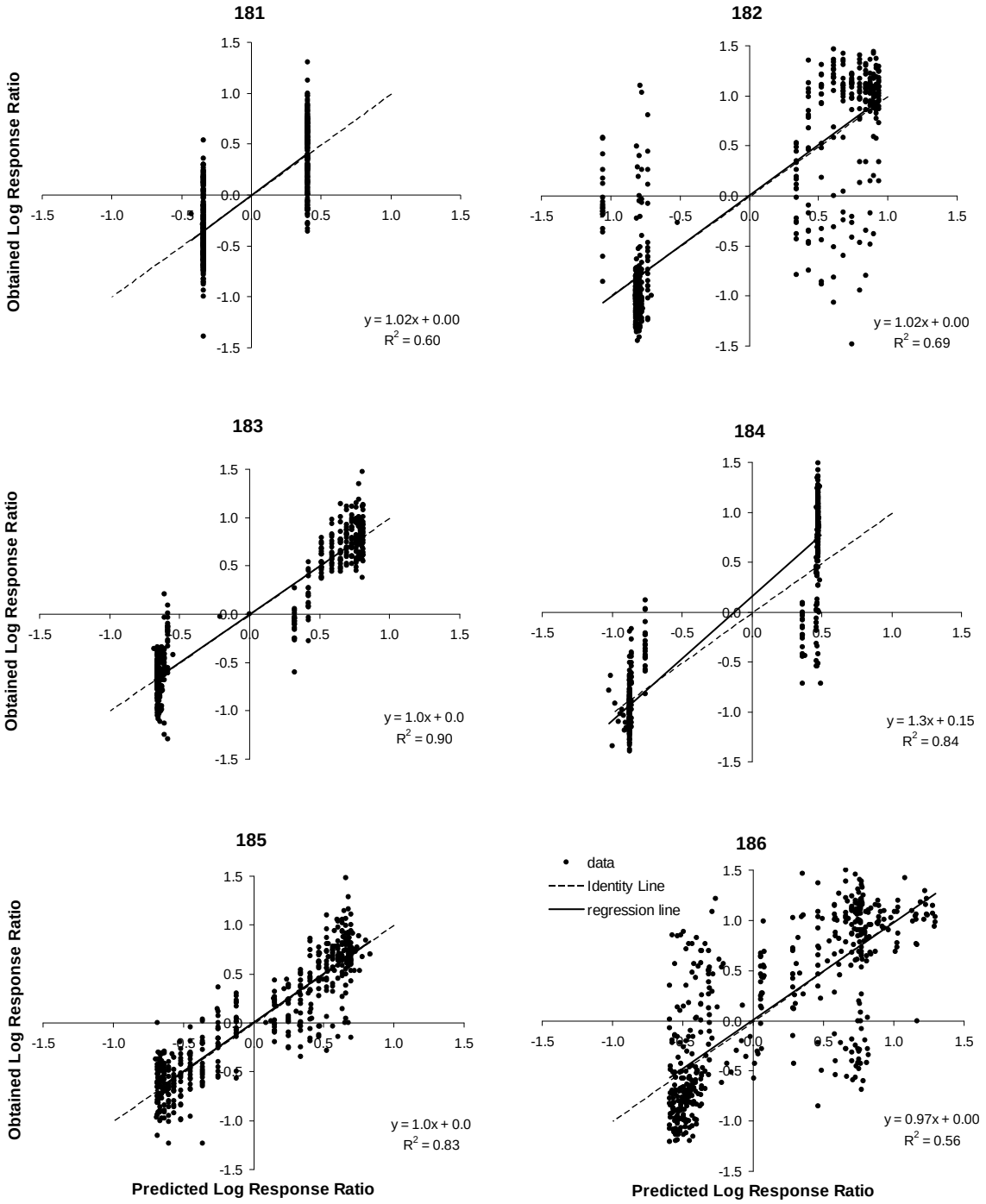


Figure 8. Individual obtained session 12th log initial-link response ratios as a function of EDM predicted log immediacy ratios averaged across replications in Phase 1. Included are the regression lines, associated best fitting r^2 values and linear regression parameters.

Figure 8 illustrates the spread of predicted vs. obtained session 12th data. The three subjects with the lowest VAC show two distinct trends in their scatter. Subject 181 appears to have two vertical lines of obtained scores suggesting that the EDM was making one of two predictions for all session 12ths suggesting the terminal-link delays were determining predicted response ratios with little effect from the initial-link. Subjects 182 and 186 appeared to have greater spread in the predicted scores, although the session 12th model fits were similar to subject 181. The lack of fit for these subjects appears to be due to variability within the data. Thus, the reduced session 12th VAC scores appear to be due to variability in the data and some inflexibility of the EDM. The remaining subjects, those with higher session 12th VAC scores, show a restricted number and also spread of EDM predictions. Therefore, EDM is flexible enough to describe the data to some degree, although there are data sets that force a constrained number of predictions which may not describe the variability in the data.

The EDM was then fitted to the Phase 2 data. As Phase 2 initial-link durations changed each session in a ascending and descending series parameter ΔS, representing carry-over between sessions, was added in the Phase 2 solver iterations. This created a total of 6 free parameters. The session 12th variance accounted for was minimised using the same process as used in Phase 1. The best fitting EDM parameters and session 12ths VAC are listed in Table 9.

Table 2.

EDM VAC and Free Parameter Values for Individual Subjects Fitted to Session 12th Data in Phase 2.

	VAC	EDM Free Parameters					
	Session 12th	σ	Δ L	Δ R	Log b	β	ΔS
181	0.84	0.33	0.33	0.33	0.08	0.00	0.24
182	0.73	0.11	0.23	0.49	0.05	1.25	0.37
183	0.76	0.13	0.36	0.23	0.05	0.13	0.00
184	0.78	0.14	0.60	0.32	0.06	1.35	0.00
185	0.77	0.20	0.33	0.24	0.00	0.50	0.00
186	0.58	0.18	0.61	1.00	-0.01	0.45	1.04
Average	0.74	0.18	0.41	0.44	0.04	0.61	0.28

Table 2 shows similar fits to Phase 2 data as Phase 1, although there is greater consistency across subjects in session 12th predicted variance accounted for by the EDM. The β parameter appears to vary independently from the ΔS parameter, suggesting that these two parameters are capturing different aspects of the data.

Figure 9 shows the session 12th obtained log initial-link response ratios as a function of EDM predictions for the individual subjects. The slopes for the best fitting regressions are shown and the group average is close to 1.0 (mean = 1.05), while average bias was 0.0 suggesting that the EDM captured the major trends in the data. Averaged across subjects, the EDM described 74% of the variance.

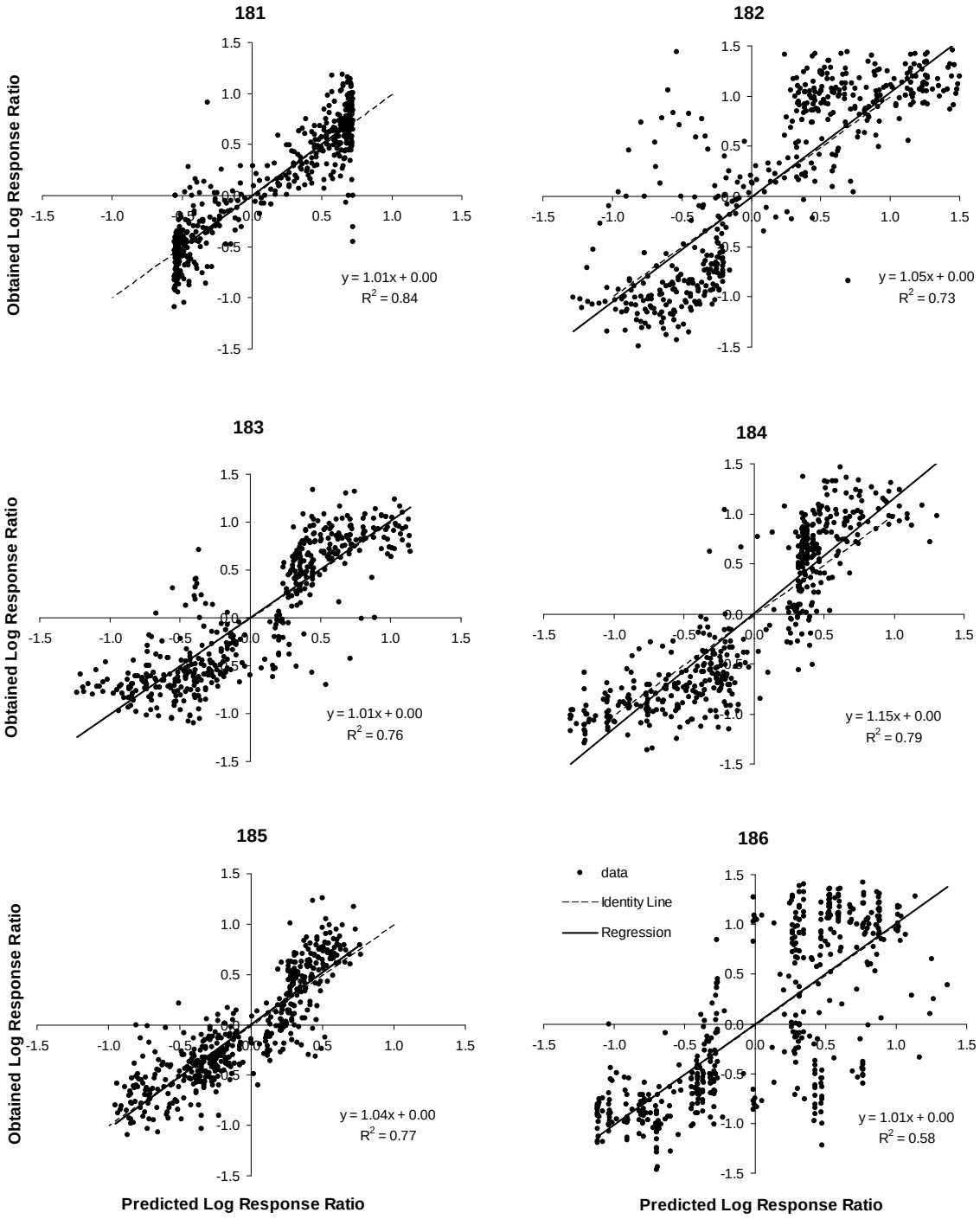


Figure 9. Individual obtained session 12th log initial-link response ratios as a function of EDM predicted log immediacy ratios averaged across replications in Phase 2. Included are the regression lines, associated best fitting r^2 values and linear regression parameters.

Figure 9 illustrates the spread of predicted vs. obtained session 12th data in Phase 2. Slopes for the best fitting regressions are shown, and the slopes are close to 1.0, suggestion the EDM captured the overall trends in the data. There is some evidence of sigmoidal curvature for some subjects, suggesting EDM predictions deviate from the obtained data. For example, Pigeons 183 and 184 appear to follow a trend that begins above the regression line and as predictions increase, falls below the regression line then rises above the regression line.

Discussion

The goal of this experiment was to explore how response allocation in a rapid acquisition concurrent-chains procedure varied when initial-link duration was changed according to an ascending and descending series. In particular, this experiment examined the relationship between response allocation and initial-link duration, and whether it was monotonic, as predicted by steady-state models for choice (Fantino, 1969; Grace, 1994; Mazur, 2001), or bitonic, as predicted by the EDM. Results showed that the preference for the FI 10 s terminal link increased for programmed initial-link durations in the range of approximately 0.01 to 7.5 s, and decreased from 7.5 s to 30 s (see Figure 6 lower panel). Thus response allocation and initial-link duration indicated a bitonic relationship contrary to existing models for steady-state choice but consistent with the extended decision model.

Although the programmed initial-link duration varied from 0.01 s to 30 s, the obtained initial-link durations were used. However, programmed rather than obtained

initial-link values were used as the independent variable to test the relationship between preference and initial-link duration, because obtained initial-link duration varied across subjects and thus was problematic as an independent variable. The upper panel of Figure 8 shows that obtained initial-link duration increased linearly with the programmed duration; however the minimum value was 8.36 s, much longer than the programmed value of 0.01 s. This is a consequence of the scheduling procedure used to equate terminal-link entries over the session: the initial link that would result in a terminal-link entry was pre-selected on each trial, any time that subjects spent responding on the other alternative after the initial-link schedule had elapsed would increase the obtained initial-link duration beyond the programmed value. Moreover, the stronger the preference in a given session, the more likely it was that subjects would be responding on the preferred alternative when a terminal link was arranged for the other schedule, which could produce an artifactual relationship between response allocation and obtained initial-link duration. However, note that the programmed duration that produced the maximum preference (7.5 s) was associated with an obtained duration of approximately 15 s (see Figure 6).

Furthermore, the group and individual regression analyses excluded the extreme initial-link durations of the ascending and descending series. This reduces the possibility that the bitonic effect was an effect of the smallest delay artificially forcing the regression analyses to find non-linear response allocations due to 0.01 being too short a duration for subjects to make a response. Moreover, for most subjects there were usually two other data points that were of a shorter duration than individual peak response allocation initial-link duration, and there was also evidence of hysteresis from both the ascending and descending series in the initial-link duration group analyses, and also bitonic functions in both series. This suggests that response

allocation from previous sessions was influencing current-session preference, and that the bitonic effect is not an artefact of aggregation across subjects.

It is important to note that the downturn in preference at short programmed initial-link durations could not have resulted from a "win stay" strategy in which entry to the richer terminal link might have been produced by a single response in the initial cycle following a reinforcer in the same terminal link. As the COD was reset at the beginning of each cycle, there was a minimum of 1 s between the first response in a cycle and that which produced terminal-link entry. To examine whether there was evidence for sequential dependency in the location of the first response in each cycle the probability of responding to the richer alternative conditional on the location of the previous terminal link was regressed against each delay and whether the previous alternative was rich or lean. The regression analyses found no significant main or interaction effects.

The EDM uses a linear-operator to calculate increments and decrements of predicted response strength. This process is dependent on whether the delay is judged short or long based on where the probability the delay is smaller than the criterion of the distribution of reinforcement history of experienced delays. In addition, the criterion also changes with changes in delay: larger delays gain relatively larger increments and lose relative smaller decrements due to the relatively less impact of a constant sigma parameter at larger delays. This was examined using the same data as the first response analysis. The first response probabilities for each subject across each delay were averaged across the four smallest delays (excluding 0.01s as this was the shortest delay so no increment was available) and the four largest for rich and lean schedules. These were chosen because they represented the two extremes in the range of delays used in this experiment. The mean increments (difference between average

values for the current less the previous delay) for both rich and lean were non-significantly (*t* test) greater for the larger delays: rich (small; 0.065, large; 0.080), and lean (small; 0.067, large; 0.068). Thus, first responses appear to be following the EDM predicted trends for both large and small delays rather than being a consequence of forcing lower response rates for short delays.

The EDM fits to the session 12th data were moderately successful. Phase 1 VAC varied between 0.56 and 0.90 and Phase 2 VAC varied between 0.58 and 0.84 suggesting that for some subjects the EDM was unable to describe the variability in the data. Nevertheless, the mostly successful fits across subjects provide additional evidence supporting the EDM predictions of the bitonic effect. Given that these results are novel and appear to challenge such well-established models as delay-reduction theory (DRT; Squires & Fantino, 1969), the contextual choice model (CCM; Grace, 1994) and the hyperbolic value-added model (HVA; Mazur, 2001), all of which have substantial empirical support, it is important to test their generality. In particular, future research needs to determine whether similar results would be obtained under steady-state conditions, with independent initial links and other terminal-link schedules. The former manipulation is especially important because independent initial links permit response allocation to go exclusively in favour of one alternative, avoiding possible ceiling effects associated with dependent scheduling.

Experiment 2

Introduction

In Experiment 1, initial-link response allocation was a bitonic function of initial-link duration when terminal schedules changed unpredictably across sessions. The aim of this experiment is to assess whether a similar result occurs when the initial-link follows an ascending and descending series and the terminal-link schedules are constant. This will test whether the bitonic effect found in Experiment 1 is evident in a procedure that has constant terminal-link values. This is an intermediate position between a quasi-random like design, such as Experiment 1, where terminal-link delays change unpredictably each session, and steady state designs where conditions remain unchanged for many sessions. Evidence for this phenomenon at an immediate level would suggest this effect is independent from the quasi-random procedure.

Method

Subjects

The same six pigeons that served in Experiment 1 were used in this experiment (numbered 181, 182, 183, 184, 185, and 186).

Apparatus

Experiment 2 uses the same apparatus as Experiment 1.

Procedure

The procedure for this experiment was similar to Experiment 1: concurrent chains procedure, sessions lasting 72 cycles or 70 mins whichever occurred first, side keys illuminated white, terminal keys assigned quasi-randomly, 1 s COD, initial-link delays were VI 10 s in phase 1 and followed an ascending and descending progression in phase 2, session running time, and terminal-link key light illumination. However, there were the following exceptions. The initial-link VI schedules contained 12 intervals constructed from an exponential progression (Fleshler & Hoffman, 1962). Separate lists of intervals were maintained for cycles in which the left or right terminal link had been selected, and were sampled without replacement so that all 12 intervals would be used three times for both the left and right terminal links each session. The 12 intervals were logarithmically spaced so that their arithmetic average was a linear function of the mean schedule duration. This created a wide distribution of intervals for each delay, with a range in the order of three times the size of each initial-link schedule..The terminal-link responses were reinforced according to FI 10 s and FI 20 s schedules, were the left and right terminal-link schedules were counterbalanced; Pigeons 181 and 182 experienced FI 20 s left terminal-links and FI 10 s right terminal-links, while Pigeons 183, 184, 185 and 186 experienced FI 10 s left terminal-links and FI 20 s right terminal-links. The training in Phase 1 (constant VI 10 s initial-links) consisted of 19, 19, 22, 19, 22 and 32 sessions for Pigeons 181 to 186, respectively. Training in phase 1 began immediately after the conclusion of Experiment 1 and phase 2 began a few days after phase 1 finished. The terminal-link schedules in phase 2 were the same as used in phase 1. The training in Phase 2 (using the same ascending and descending initial-links as in Experiment 1) consisted of 49,

50, 52, 47, 51 and 34 sessions for Pigeons 181 to 186, respectively. Subjects were started at a random initial-link value and delay.

Results

Figure 10 illustrates the relationship between obtained and programmed initial-link time: the average obtained time in the initial links (averaged across subjects) was plotted against the programmed initial-link duration (upper panel). In addition, the average log response ratio for each initial-link delay across subjects was plotted to illustrate the effect of the changing initial-link (bottom panel).

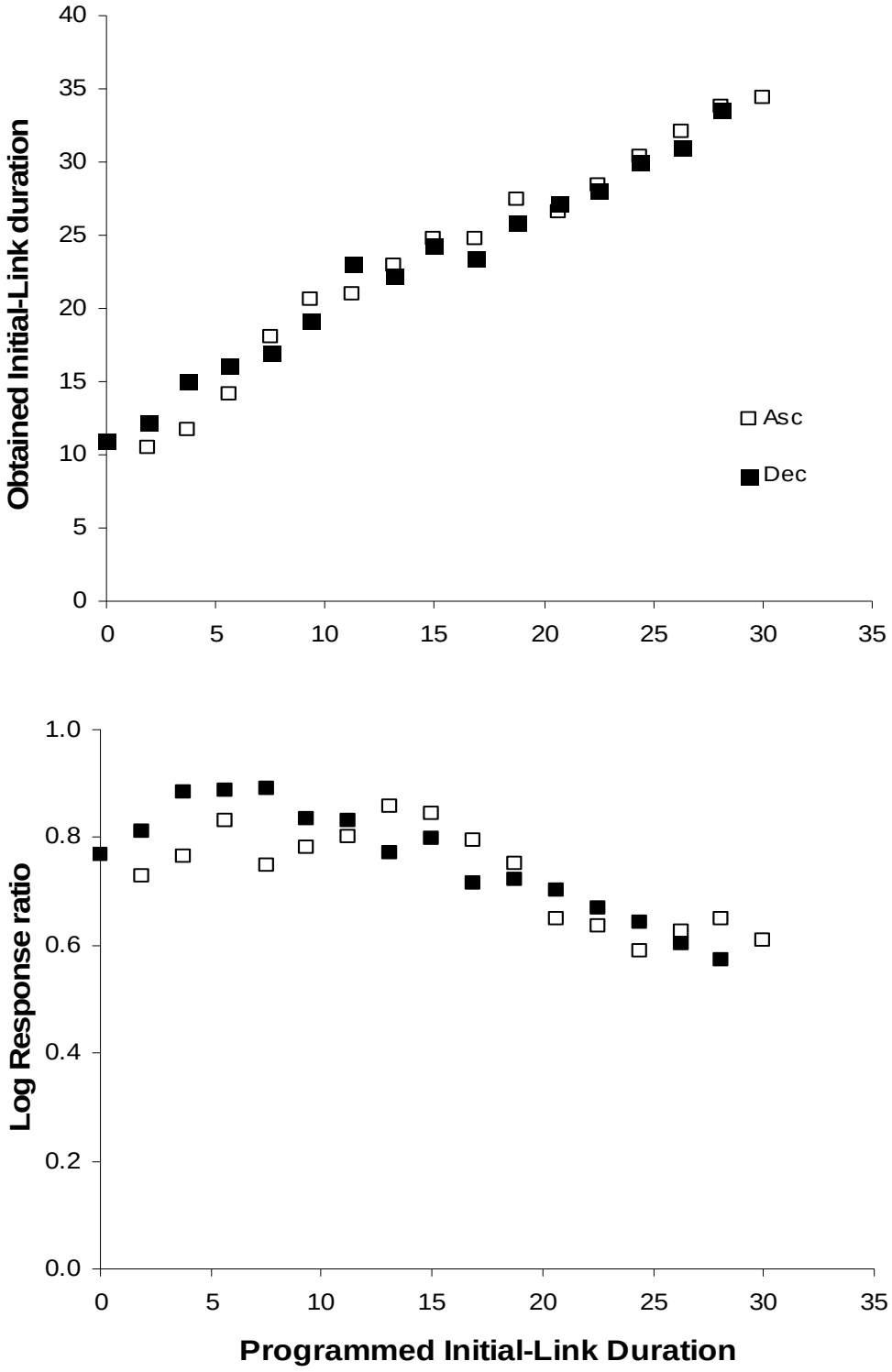


Figure 10. Group average obtained initial-link duration as a function of programmed initial-link duration (top panel) for ascending and descending series. Group average log response ratios as a function of programmed initial-link duration (bottom panel). The bottom panel illustrates the bitonic effect for both ascending and descending series.

The upper panel of Figure 10 shows that the obtained initial-link duration increases linearly with programmed duration: the equation $y = 0.82x + 10.71$ accounted for over 96% of the variance. A repeated-measures ANOVA with initial-link duration (0.01s to 30s) and series (ascending/descending) as factors confirmed the effect of programmed initial-link value, $F(14,70) = 63.63$, $p<0.001$, with no significant interactions. Figure 10 shows minimum obtained initial-link duration of about 10.71s. This occurred because the dependent scheduling arrangement, like the previous experiment, required responses on the non-preferred alternative. The consequence of this arrangement was to lengthen the average time spent in the initial-link for the shorter initial-link durations.

For each subject, log response ratios were averaged across replications of the ascending and descending series for each initial-link schedule. Thus, for each subject, an average log response ratio at each delay for the ascending and also the descending series was calculated. The bottom panel in Figure 10 shows the average log response allocation for each series as a function of programmed initial-link duration (the sign of the log response ratio was reversed for the two pigeons with FI 20 FI 10 terminal links prior to averaging). Figure 10 shows response allocation was a bitonic function of initial-link duration for both series. A repeated-measures ANOVA with initial-link delay and series as factors found a significant effect of initial-link delay, $F(14,70) = 7.96$, $p<0.001$ and a significant interaction between series and initial-link delay, $F(14,70) = 2.50$, $p<0.05$. The main effect of series was not significant. Planned polynomial contrasts found significant linear, quadratic, and cubic trends for initial-link duration, $F(1,5) = 19.01$, $p<0.05$, $F(1,5) = 10.85$, $p<0.05$, $F(1,5) = 6.63$, $p<0.05$, respectively. The significant linear component confirms the "initial-link effect", where response allocation becomes less extreme as initial-link duration increases. The

significant quadratic and cubic trends found in the statistical analysis show that response allocation was non-monotonic. Figure 10 shows that log response allocation increased during the smaller initial-link durations but then fell as initial-link duration continued to increase. The cubic component emerged as the ascending series appeared to reach a minimum at approximately 25 s and then increased at longer durations, whereas the descending series continued to decrease until the end of the series.

The significant interaction between initial-link duration and series occurred because the ascending series typically had higher log response ratios than the descending series for long initial-link delays, whereas the descending series had higher response ratios for shorter delays. This indicates that log response ratios shifted to the right for the ascending series compared to the descending series.

Figure 11 shows the individual subjects' log response allocation as a function of initial-link duration. For each pigeon there appears by visual inspection to be a non-monotonic trend for both series. This is similar to the average data: an early increase in log response allocation at short initial-link durations, and a later decrease as initial-link delay continues to increase. As expected, there is greater variability in the individual data compared to the group data, with a possible bitonic effect more clearly evident for some subjects than for others. For example, data for pigeons 182 and 183 show the effect, whereas log response allocation for pigeon 184 does not change systematically with initial-link duration. However, all pigeons except 184 show the upturn in preference at short initial-link durations and a gradual decrease at longer ones, suggesting that the bitonic effect was generally obtained in the individual data and was not an artefact of the averaging process. Note that the degree in upturn in preference varied across pigeons, with some showing a stronger effect than others.

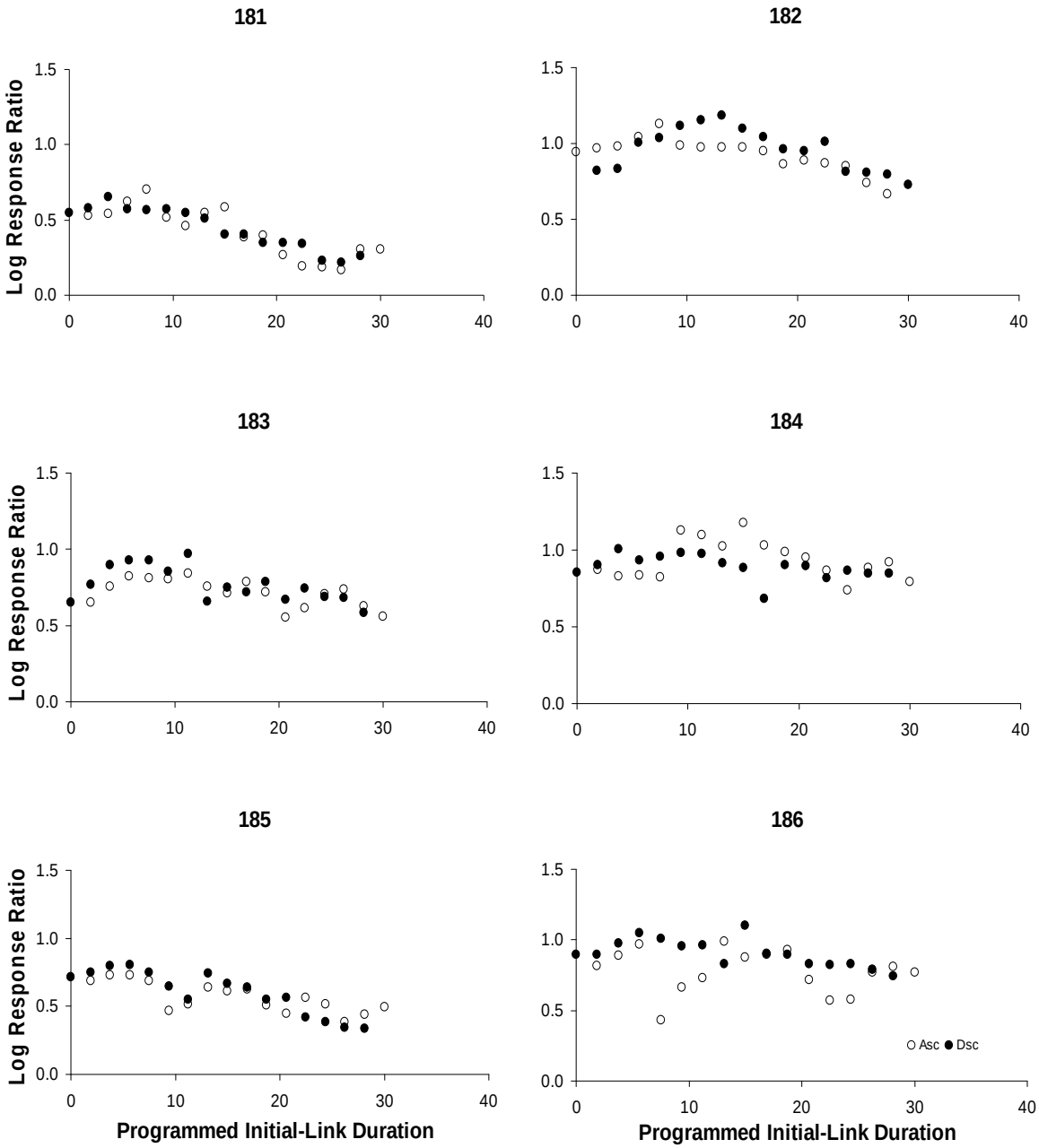


Figure 11. Individual subject log response ratios as a function of programmed initial-link duration for ascending and descending series. The filled symbols is the average log response ratio for the descending series at each delay while the unfilled symbols is the average log response ratio for each delay in the ascending series.

Discussion

Experiment 2 was designed to investigate how response allocation changed when terminal links were constant and initial-link duration changed according to an ascending and descending series. Specifically, this experiment examined whether the relationship of response allocation and initial-link duration was monotonic, as predicted by models for steady-state choice (Fantino, 1969; Grace, 1994; Mazur, 2001), or bitonic as predicted by the EDM. The group results showed preference for the FI 10 terminal link, increased for programmed initial-link durations from approximately 0.01 to 7.5 s, and decreased from 7.5 to 30 s (see Figure 6 lower panel). Repeated-measures ANOVA also showed statistical significant deviations for linear, quadratic and cubic components in the data. However, the bitonic effect was less distinct in the individual analyses, with noticeable deviations from bitonicity across most subjects. The difference between experiments is likely due to the difference in experimental procedures and number of training sessions. Experiment 1 had more numerous phase 1 and 2 sessions than Experiment 2. Moreover, the PRBS design is likely to sensitise subjects to the current session more so than a counterbalanced terminal-link. The combination of these two issues would disrupt the precise discriminations in Experiment 1, although subject 181 showed a clear bitonic effect in the ascending series.

Experiment 3

Introduction

Experiments 1 and 2 showed that group preference in concurrent chains was a bitonic function of initial-link duration when the initial links were changed systematically across sessions according to an ascending and descending series. This result is consistent with the bitonic prediction of the EDM. Experiment 3 was designed to investigate the effects of initial-link duration on preference in a rapid acquisition procedure using short and long initial-links. In different conditions, the initial links were either relatively short (VI 5 s) or long (VI 30 s), while the left terminal-link duration was always FI 8 s and the right changed unpredictably across sessions according to a quasi-random sequence binary sequence.

The rationale for this experiment was to study the effect of short and long initial-link durations in a rapid-acquisition design where the initial-link schedule was constant across blocks of sessions and relative terminal-link immediacy changed in a quasi-random binary sequence. Although the design precludes the possibility of testing for a bitonic effect, I expected to be able to show that sensitivity to the current-session immediacy ratio was less when the initial links were long than when they were short. In addition, within session analyses investigating the effect of changing initial-link and terminal-link delay was conducted to examine whether there were differential effects during acquisition.

Method

Subjects

Four pigeons of mixed breed, numbered 195, 196, 197 and 198, served as subjects and were maintained at 85% of their free-feeding weight plus or minus 15 g through appropriate post-session feeding. Subjects were housed individually in a vivarium with a 12h:12h light/dark cycle (lights on at 0600), with water and grit freely available in the home cages. All pigeons were experienced with a variety of experimental procedures.

Apparatus

Experiment 3 uses the same apparatus as Experiments 1 and 2.

Procedure

The procedure was similar to Experiment 1: concurrent-chains procedure, sessions lasting 72 cycles or 70 mins whichever occurred first, side keys illuminated white, terminal keys assigned quasi-randomly, 1 s COD, and terminal-link key light illumination. However, there were the following exceptions. The experiment consisted of two conditions, which varied only in terms of the initial-link schedule. In the 'short' condition the initial-link schedule was VI 5 s, while in the 'long' condition the initial-link schedule was VI 30 s. The sequence of the conditions was counter balanced across pigeons, with the sequence short-long-short-long for pigeons 195 and 196, and long-short-long-short for pigeons 197 and 198. The initial-link VI schedules contained 12 intervals constructed from an exponential progression (Fleshler &

Hoffman, 1962). Separate lists of intervals were maintained for cycles in which the left or right terminal link had been selected, and were sampled without replacement so that all 12 intervals would be used three times for both the left and right terminal links each session. The 12 intervals for each delay were logarithmically spaced so that their arithmetic average was a linear function of the mean schedule duration. This created a wide distribution of intervals, with a range in the order of three times the size of each initial-link schedule. For both conditions the left terminal-link alternative was always 8 s whilst the right alternative changed each session in a quasi-random 31-step binary sequence (either 4 s or 16 s). Conditions were terminated when the regression analyses of the log response allocation to the initial link showed that response allocation during the last 20 sessions showed strong sensitivity to the immediacy ratio in the current session and negligible position bias. The last 20 sessions were chosen because some conditions had less than 30 available sessions with three previous lags (that is, for some subjects, there was no lag 3 available at session 30 due to missing data for some subjects due to computer malfunctions and human error). Thus, in the interests of consistency across subjects, the last 20 sessions for each subject were analysed rather than analysing different numbers of sessions across subjects. Subjects experienced the same number of condition replications, although the number of sessions was different for each condition. The analyses used the last three conditions. The number of sessions in order was 93, 62 and 43 for each subject.

Results

To assess the relationship between response allocation and the immediacy ratios in the current and prior sessions, we used a generalized-matching model:

$$\log\frac{B_{0L}}{B_{0R}} = a_0 \log\frac{1/D_{0L}}{1/D_{0R}} + a_1 \log\frac{1/D_{1L}}{1/D_{1R}} + a_2 \log\frac{1/D_{2L}}{1/D_{2R}} + \ldots + \log b \quad (28)$$

where B and *D* refer to initial-link response rate and terminal-link delay, respectively, subscripted for choice alternative (L and R) and lag (0 through 3; 0 = lag 0). The parameters $a_0 \ldots a_4$ quantify sensitivity to reinforcer immediacy (i.e., reciprocal of delay) at each lag, and log *b* is a bias parameter. Equation 28 was applied to the data for individual subjects from the last 20 sessions from the last three conditions were analysed. The regression statistics from equation 28 are shown in Table 3.

Table 3.

Multiple Regression Statistics: Condition and Lag0-3 Sensitivity Values

	195				196			
		Long	Short	Long		Long	Short	Long
	Lag0	**0.75**	**1.80**	**1.25**	Lag0	**0.80**	**1.00**	**0.70**
	Lag1	-0.01	-0.02	-0.01	Lag1	0.03	-0.07	0.06
	Lag2	0.10	0.01	0.03	Lag2	0.10	-0.07	0.06
	Lag3	-0.05	0.00	**0.04**	Lag3	0.10	-0.08	0.03
R^2		0.85	0.98	0.98		0.45	0.74	0.62
	197	Short	Long	Short	198	Short	Long	Short
	Lag0	**1.30**	**0.60**	**1.40**	Lag0	**0.70**	**0.32**	**0.70**
	Lag1	**0.40**	0.07	0.04	Lag1	**0.24**	0.08	0.07
	Lag2	**0.40**	0.13	0.07	Lag2	**0.24**	0.03	-0.03
	Lag3	**0.35**	0.02	0.03	Lag3	0.00	0.22	-0.40
R^2		0.94	0.88	0.96		0.77	0.11	0.75

Bold $p<0.05$

Table 3 shows multiple regression sensitivity values for each Pigeon (the log immediacy for lags 0 to 3 is regressed against the current session's log response allocation), response allocation was positive and statistically sensitive to Lag 0 from the first condition to the third condition. Additionally, Pigeon 195's response allocation was also sensitive to lag 3 in the third condition. For two subjects, Pigeons 197 and 198, response allocation was sensitive to lags 1 and 2 in the first condition, whilst Pigeon 197 was also sensitive to lag 3 in the first condition.

Figure 12 illustrates the effect of initial-link duration on current session sensitivity. For each subject, current session sensitivity was greater for the short than the long condition for both replications. This demonstrates the initial-link effect, in which response allocation is more extreme with short than with long initial-link duration (Fantino, 1969) using a quasi-random procedure.

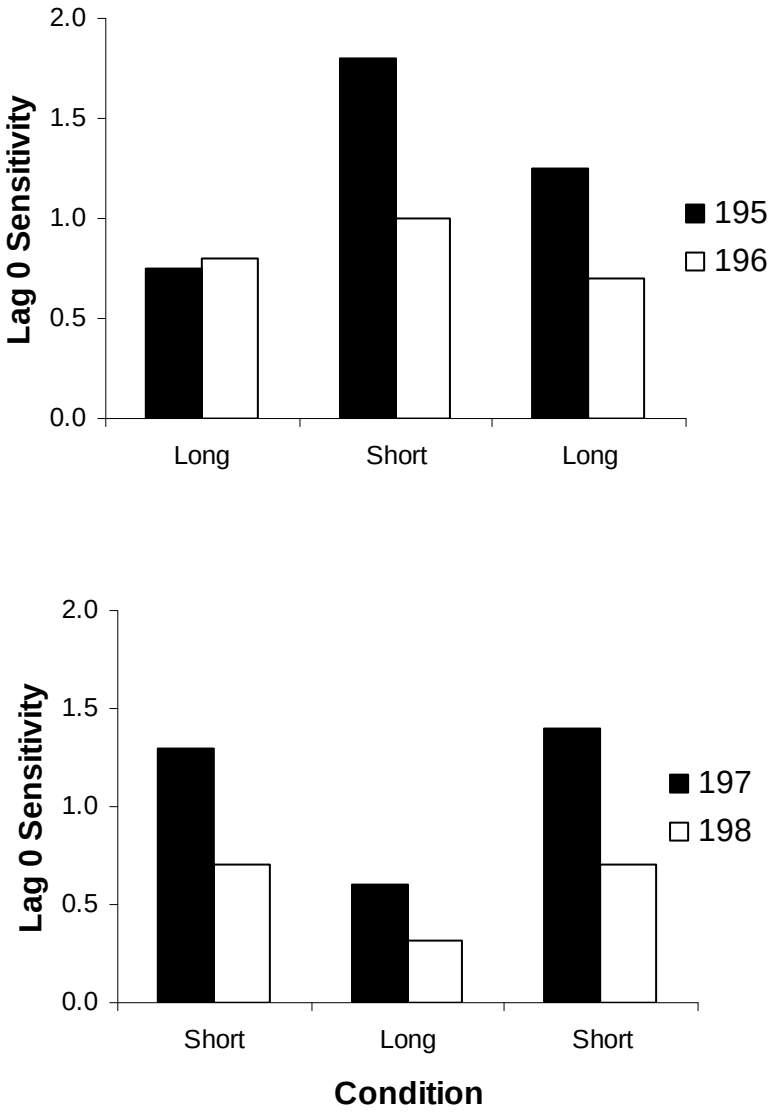


Figure 12. Individual subject current session sensitivity for each condition. Pigeons 195 and 196 experienced a long-short-long order (top panel), while pigeons 197 and 198 experienced a short-long-short order (bottom panel). In all cases, the short initial-link duration had higher within-subject current session sensitivity values irrespective of order.

Figure 12 illustrates that although all subjects showed sensitivity to the changing initial-link duration there were individual differences across subjects. For example, the changes in sensitivities for Pigeon 196 were modest whilst the changes in sensitivities for Pigeons 195 and 197 were more obvious. These differences appear to be independent from their starting values: Pigeon 198 had a relatively low beginning sensitivity value but showed distinct differences in later conditions. Another interesting result was that for three of the four pigeons (196, 197, and 198) sensitivities appeared to return to their starting values in condition one when replicated in the third condition.

To examine the course of acquisition within sessions, response allocations for the last 20 sessions were plotted for each session 12^{th} (the total number of trials in each session was 72, making 6 trials per session 12^{th}). Conditions were averaged separately depending on whether the initial-link was short or long and whether the right terminal-link schedule was FI 4 s or FI 16 s. This created four conditions for each subject: short initial-link and short terminal-link, short-initial-link and long terminal link, long initial-link and short terminal-link, and long initial-link and long terminal-link. Figure 13 illustrates the acquisition curves. For three pigeons, log response ratios appear to separate based on the terminal-link duration indicating the terminal-link dominating response allocation as suggested by the terminal-link effect, while the most extreme response allocations were mostly for short than long initial-link delays indicating the initial-link effect. However for pigeon 198 there appears to be a shift in response bias, such that preference is much stronger for the left key in the short compared to long condition. Reasons for this shift in bias are unclear.

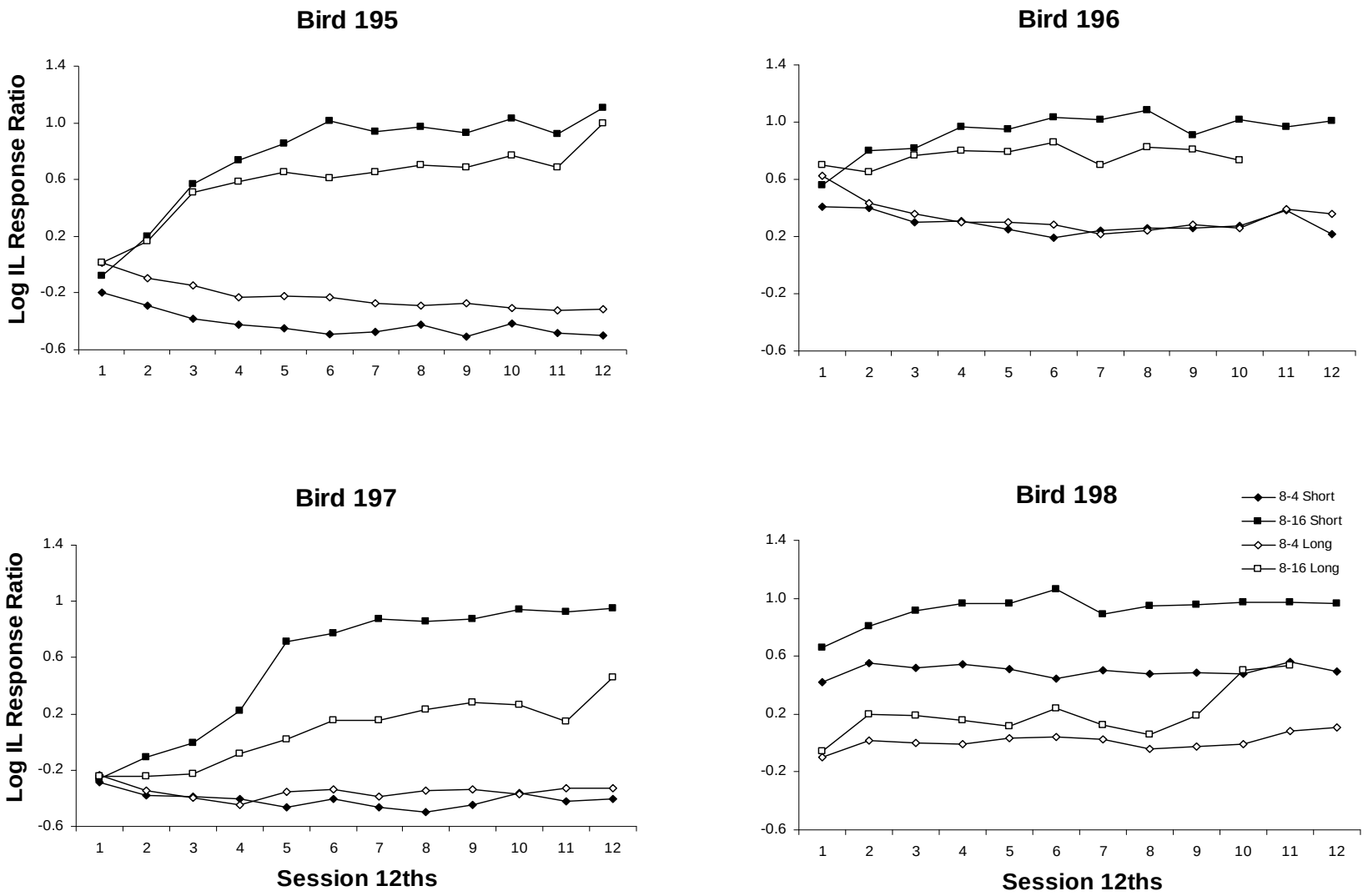


Figure 13. Individual subject log initial-link response ratio as a function of session 12ths and terminal-link duration pair

Discussion

This experiment was designed to explore how response allocation in a PRBS concurrent-chains procedure varied when blocks of initial-link durations were short or long. For all pigeons, response allocation was found to be sensitive to the current session from the first condition. Moreover, the higher lag sensitivities were more likely to be statistically significant. This replicates previous findings using the PRBS-like design in concurrent-schedules and concurrent-chains (Grace, Bragason, & McLean, 2003; Schofield & Davison, 1997). However, this experiment differs from the Grace et al. experiment where initial-links were VI 10 s and the left terminal-link changed in a PRBS design to either 4 s or 16 s. In the present study both initial- and terminal-link durations were varied: the initial-link duration changed in blocks of short VI 5 s and long VI 30 s, while the terminal-link used a quasi-random binary sequence design where the right link was always 8 s and the left changed unpredictably to either 4 s or 16 s. Nevertheless, average values of the current session sensitivity coefficients were 1.2 and 0.75 in the short and long conditions, respectively. This was similar to the average 1.0 sensitivity coefficient reported by Grace et al.'s (2003).

When short and long initial-link durations were compared in this experiment, sensitivity was higher for all subjects for the short initial-link duration. This is analogous to the initial-link effect, where higher preference occurs for the preferred alternative for short compared to long initial-link durations (Fantino, 1969). These results suggest that sensitivity to the current session's immediacy ratios in rapid-acquisition concurrent chains is affected by initial-link duration in a manner similar to traditional steady state experiments.

The within-session response allocations showed three out of four subjects' acquisition curves were controlled by both the terminal-link and the initial-link ratio, although the initial-link appears dominate as response ratios were more extreme for shorter initial-links even when the terminal delays were large. One possible reason springs to mind: the presentation of the initial-link was more consistent: subjects experienced blocks of short or long initial-link conditions for around 30 sessions whilst the terminal-link changed unpredictably across sessions. The terminal-link durations also showed an interesting result: the long durations had higher response allocations than the short, suggesting the terminal-link effect in a quasi-random binary sequence.

Separate effects of initial and terminal links have been found in research that examines initial- and terminal-link responding. For example, Grace, Berg and Kyonka (2006) explored whether performance in terminal-link no-food trials was related to initial-link performance in a concurrent-chains procedure. They found preference for the shorter terminal-link alternative was higher with short initial links, replicating the initial-link effect. However, response allocation in the no-food peak procedure was unaffected by initial-link duration but approximated the terminal-link duration. Local measures of analyses supported this finding: increases in the initial-link was associated with greater number of visits to the non-preferred alternative, whilst regression analyses found no additional evidence of the initial-link contributing to increases in performance beyond predictions based on terminal-link duration. In conclusion, these results appear to suggest that the initial-link effect was obtained when reinforcer contingencies changed each session according to a quasi-random binary sequence design, and the blocks of short and long initial-link conditions dominated within session allocation.

Chapter Three: Terminal-Link Experiments

Experiment 4

Introduction

Although the primary purpose of the present research was to examine the predictions of the EDM, in this experiment another linear-operator model of choice acquisition, LINOP (Grace, 2002a), was chosen as a comparator as it has similarities and also important differences to the EDM for a situation that is intermediate between steady-state designs and the PRBS procedure. Specifically, this experiment investigates the changes in response allocation when the relative terminal-link reinforcer immediacies followed a systematic ascending and descending series. This procedure is an intermediate situation because the terminal-link immediacy ratio changes every session, but the changes are correlated because they follow a predictable pattern. In this experiment, the left terminal link was always FI 8 s, while the schedule value for the right terminal link changed from 2 s to 32 s and back to 2 s (or from 32 s to 2 s to 32 s) through a geometrically-spaced 17-step series. Figure 14 show predictions for this situation by the EDM (left panel) and LINOP (right panel).

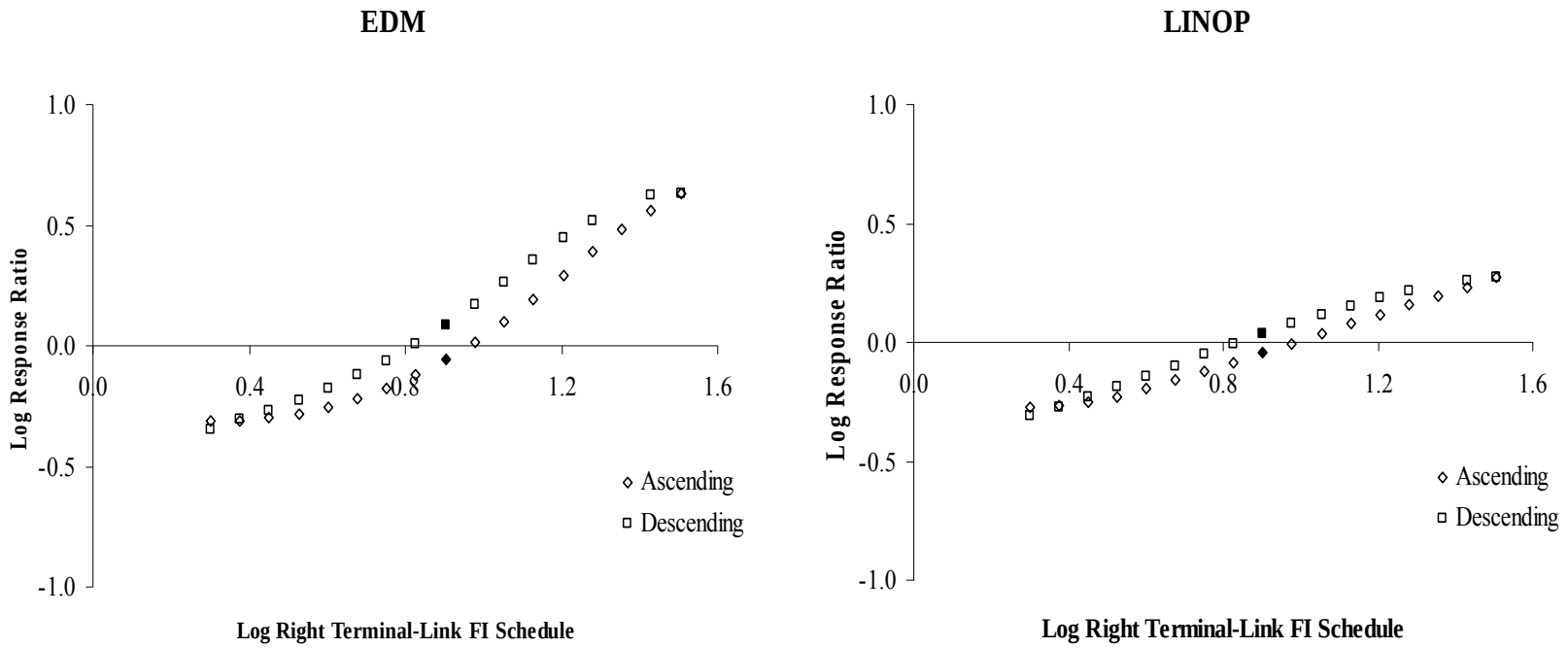


Figure 14. Log initial-link response ratios as a function of log terminal-link immediacy ratios predicted by EDM and LINOP. Unfilled symbols are predicted log response ratios for the ascending and descending series. Filled symbols are the predictions at the mid point for each series.

In both panels, the predicted log initial-link response ratio is plotted as a function of the log FI schedule value for the right terminal link. Predictions are shown for the 15 values between 2 s and 32 s, which were arranged during both the descending and ascending series. Predictions depend on the specific parameter values used, but the qualitative trends evident in Figure 14 are robust. Both models predict that preference for the left terminal link (FI 8 s) is overall greater on the descending than ascending series. This would correspond to a hysteresis effect; at the start of the descending series, the right terminal link from the previous session is FI 32 s, and so a nearly maximal preference for the left alternative should have been reached. However, the models differ in terms of the strength of preference for the shorter terminal link depending on whether the right terminal link is less than or greater than 8 s. The filled symbols in Figure 14 indicate when the right terminal link was 8 s, and divide both series into halves in which the absolute values of the log immediacy ratios are equal. According to the LINOP model, the strength of preference for the left alternative when the right delay is greater than 8 s (points to the right of the filled symbols) is the same as the strength of preference for the right alternative when the right delay is less than 8 s (points to the left of the filled symbols). However, the EDM predicts that the strength of preference should be overall greater when the right delay is longer than 8s. This exemplifies the terminal-link effect (Grace, 2004; Grace & Bragason, 2004; MacEwen, 1972) that preference should be more extreme with overall longer delays, with the delay ratio held constant.

Method

Subjects

Eight pigeons of mixed breed, numbered 221, 222, 223, 224, 191, 192, 193 and 194 served as subjects and were maintained at 85% of their free-feeding weight plus or minus 15 g through appropriate post-session feeding. Subjects were housed individually in a vivarium with a 12h:12h light/dark cycle (lights on at 0600), with water and grit freely available in the home cages. Pigeons 221, 222, 223 and 224 (Group Experienced) were experienced with the rapid-acquisition concurrent-chains procedures and had served as subjects just prior to the start of the present study, whereas pigeons 191, 192, 193 and 194 (Group Naïve), although experienced with other procedures, had no prior training with rapid-acquisition concurrent chains.

Apparatus

Four standard three-key operant chambers, 32 cm deep x 34 cm wide x 34 cm high, were used. The keys were 21 cm above the floor and arranged in a row. In each chamber there was a houselight located above the centre key that provided general illumination, and a grain magazine with an aperture centred 6 cm above the floor. The magazine was illuminated when wheat was made available. A force of approximately 0.15 N was necessary to operate each key. Each chamber was enclosed in a sound-attenuating box, and ventilation and white noise were provided by an attached fan. Experimental events were controlled and data recorded through a microcomputer and MEDPC® interface located in an adjacent room.

Procedure

For all pigeons, training started immediately in a concurrent-chains procedure. The houselight provided general illumination at all times except during reinforcer delivery. With few exceptions, sessions were run daily and at approximately the same time (1000h for Group Experienced; 1200h for Group Naive).

Sessions ended after 72 initial- and terminal-link cycles or 70 min, whichever occurred first. At the start of a cycle, the side keys were illuminated white to signal the initial links. An entry was assigned pseudo-randomly to the left or right terminal link with the constraint that in every 6 cycles, 3 entries occurred to each terminal link. An initial-link response produced an entry into a terminal link provided that: (a) it was made to the preselected key; (b) an interval selected from a VI 10-s schedule had timed out; and (c) a 1-s changeover delay (COD) was satisfied— at least 1 s had elapsed following a changeover to the side for which terminal-link entry was arranged.

The VI 10-s initial-link schedule did not begin timing until the first response had occurred in each cycle, to allow any pausing after the completion of the previous terminal link to be excluded from initial-link time. The VI 10-s schedule contained 12 intervals constructed from an exponential progression (Fleshler & Hoffman, 1962). Separate lists of intervals were maintained for cycles in which the left or right terminal link had been selected, and were sampled without replacement so that all 12 intervals would be used three times for both the left and right terminal links each session.

When a terminal link was entered, the colour of the side key was changed (left key to red, right key to green) while the other key was darkened. Terminal-link responses were reinforced according to FI schedules. When a response was reinforced

all lights in the chamber were extinguished, and the grain magazine raised and illuminated for 3 s.

The FI schedule value for red (left) terminal link was always 8 s, and the value for the green (right) terminal link was one of the following: 2, 2.38, 2.83, 3.36, 4, 4.76, 5.66, 6.73, 8, 9.51, 11.31, 13.45, 16, 19.03, 22.63, 26.91 or 32 s. The right terminal-link schedule values were equally spaced in logarithmic terms, and occurred in an ascending or descending series across sessions. For example, the 2 s delay was always followed by 2.38 s in the next session, and 2.83 s in the session after that (like in the order listed above), whereas the 32 s delay was always followed by delays in the reverse order (for example, 26.91 s in the next session, 22.63 s, in the following session, etc.).

For two pigeons in Group Experienced (221 and 222), the right terminal link began at 2 s and three ascending and descending series were completed; for pigeons 223 and 224, the right terminal link began at 32 s and three descending and ascending series were completed. All pigeons in Group Naïve first received 21 sessions in which both terminal-link schedules were FI 8 s. The purpose of this training was to establish a baseline from which the effects of the ascending and descending series could be assessed. Delays were then increased across sessions to 32 s for pigeons 191 and 192 according to the geometric series, and decreased across sessions to 2 s for pigeons 193 and 194. All pigeons then completed three descending and ascending series (191 and 192) or ascending and descending series (193 and 194).

Results

Figure 15 shows response allocation and the programmed immediacy ratio plotted over sessions for all subjects across the three ascending and descending series. Figure 16 illustrates that response allocation for all subjects in both groups tracked changes in the immediacy ratio. Response allocation increasingly favoured the left initial link during the ascending series (in which the right terminal link changed from 2 s to 32 s), and the right initial link during the descending series (in which the right terminal link changed from 32 s to 2 s). Individual differences are also apparent. For example, shifts in response allocation were small and gradual across sessions for some pigeons (for example, 222, 223, 224 in Group Experienced, and 192 in Group Naïve), corresponding to changes in the log immediacy ratio, but large changes were evident for others (such as pigeons 221 and 193). There was a pronounced bias toward the left initial link for pigeon 223, and to a lesser extent for pigeons 193 and 194. Overall, there appears to be no systematic difference between Group Experienced and Naïve in terms of changes in response allocation across sessions.

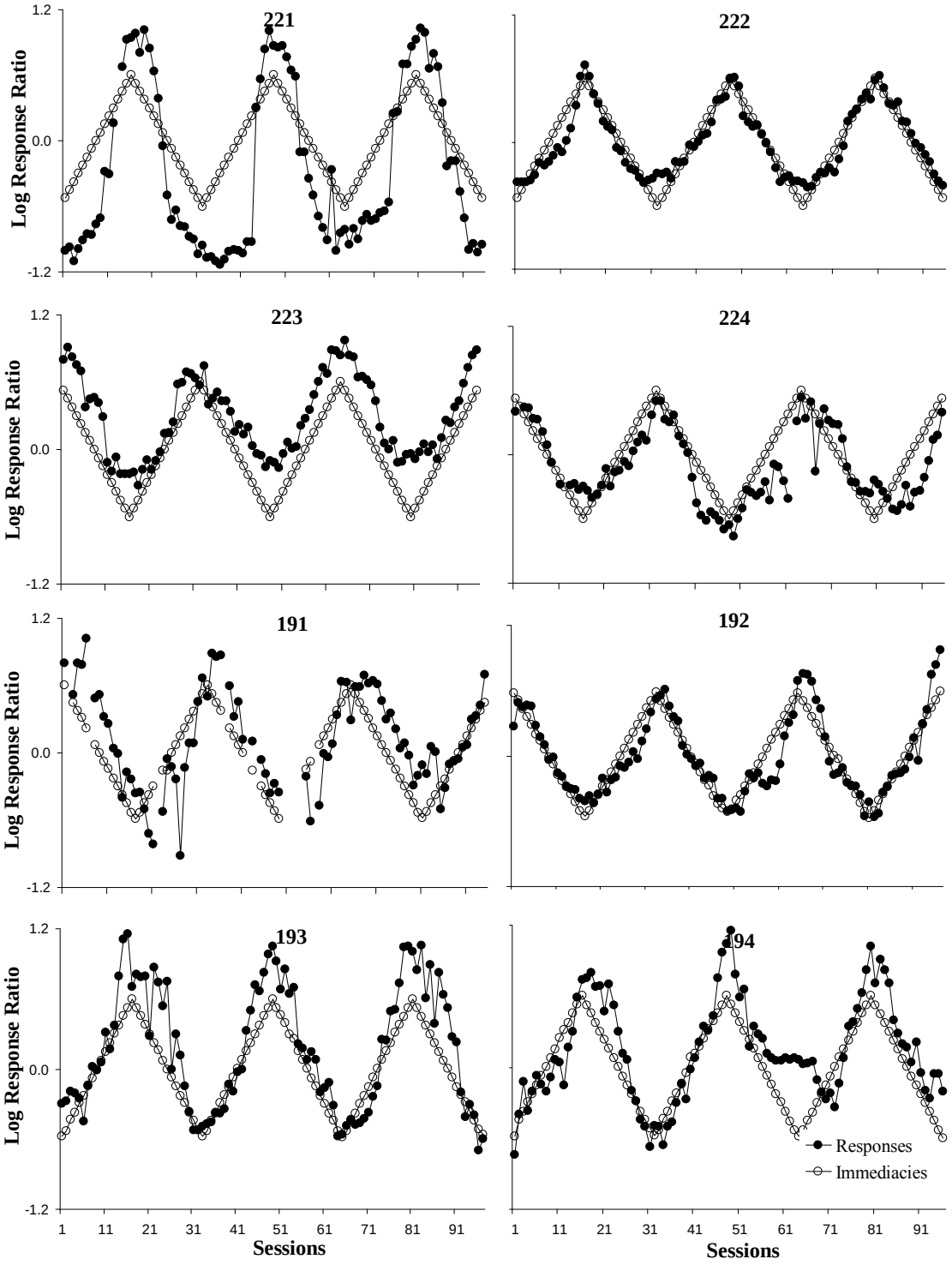


Figure 15. Obtained log initial-link response allocation and log programmed terminal-link immediacy ratios across all three replications for the ascending/descending series for subjects in Group Experienced (Pigeons 221, 222, 223, and 224) and Naïve (Pigeons 191, 192, 193, and 194).

To assess results in Figure 15 more systematically, individual-subject data were entered into a repeated-measures analysis of variance (ANOVA) with group (Naïve or Experienced) as a between-subjects factor and log immediacy ratio, replication (1^{st}, 2^{nd} or 3^{rd} presentation of a series) and series (ascending or descending) as within-subjects factors (the missing data points were due to computer malfunctions on those days). The main effects of series (ascending versus descending) and log immediacy ratio were significant, $F(1,6) = 28.96$ and $F(14,84) = 42.73$, both $p < 0.01$, respectively, whereas those of group and replication were not, $F(1,6) = 1.23$ and $F(2,12) = 3.02$, both *ns*.

There were two significant interactions, replication x log immediacy ratio, $F(28,168) = 1.55$, $p < 0.05$, and series x log immediacy ratio, $F(14,84) = 4.39$, $p < 0.01$. Analysis of simple effects showed that response allocation favoured the right initial link relatively more during the second replication when the delay was 8 s and 9.51 s, and favoured the left initial link relatively more during the third replication when the delay was 16 s and 22.63 s. Although reasons for these differences are unclear, the effects were small and apparently unsystematic in the context of the overall changes in preference.

To highlight the series x log immediacy interaction, Figure 16 shows log response ratio as a function of the log terminal-link immediacy ratio, averaged across replications. All subjects responded relatively more to the left initial link during the descending series, especially for immediacy ratios in the middle of the range, but response allocation to converged at the most extreme immediacy ratios.

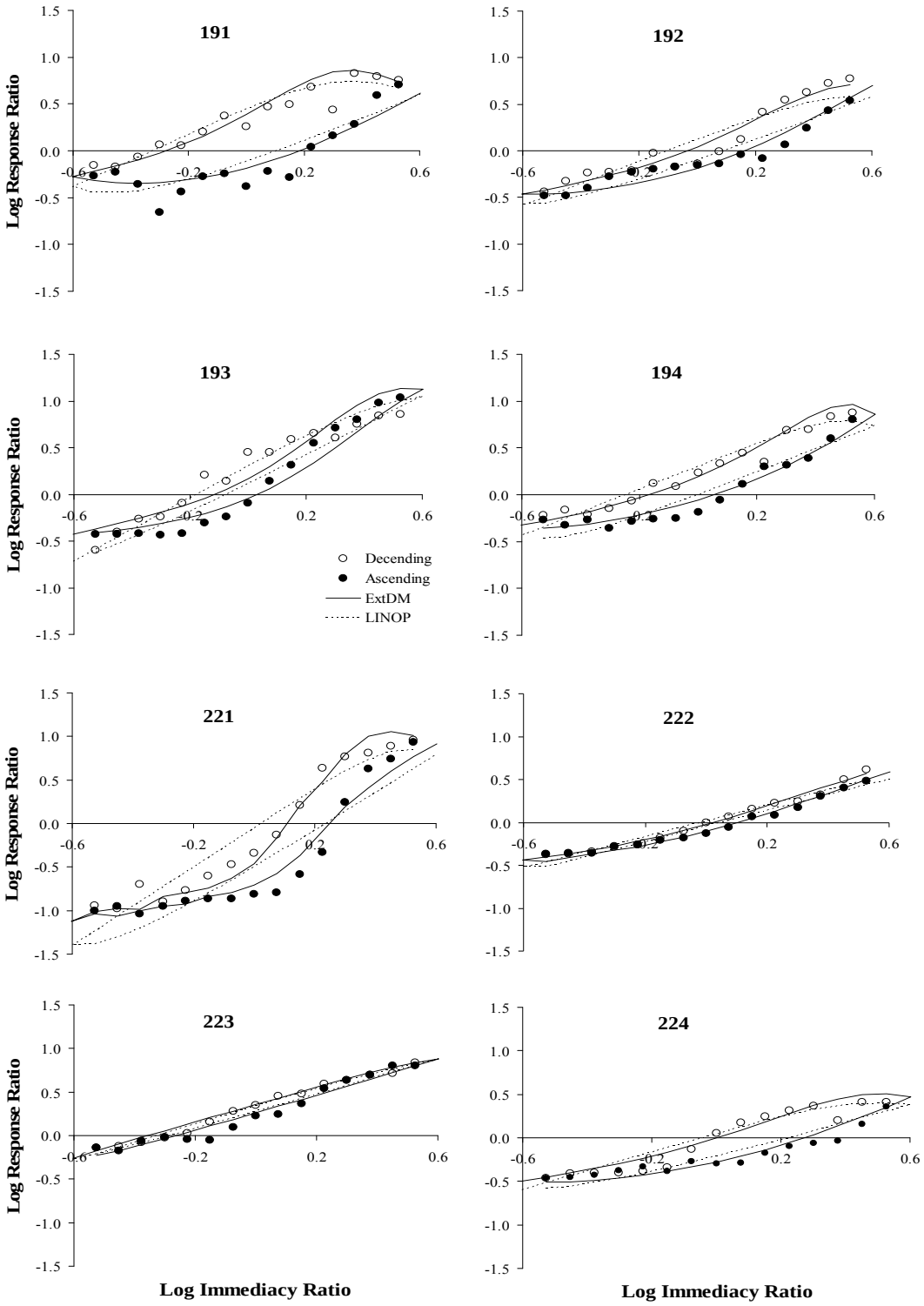


Figure 16. Obtained log initial-link response ratios as a function of programmed log terminal-link immediacy ratios, for both ascending and descending series, averaged across replications for individual subjects. Predictions of EDM and LINOP are also shown by solid and dashed lines, respectively.

Overall, the pattern might be described as a "bubble" near the middle of the immediacy ratio range, and indicates a hysteresis effect. This effect occurred as follows: at the end of the ascending series, the right terminal-link delay was 32 s and response allocation strongly favoured the left key. The preference for the left key persisted while the right-key delay decreased during the descending series, but eventually responding switched to favour the right when the delay became sufficiently short. When the delay was 2 s at the end of the descending series, response allocation strongly favoured the right key. As the delay began to increase in the ascending series, preference for the right key persisted until the delay became sufficiently long, when it switched to the left key. Thus, the persistence in response allocation at the end of both series produced an overall increased preference for the left key in the descending series, creating the bubble pattern. The magnitude of this effect varied across subjects; it was strong for pigeons 221 and 191, but relatively weak for pigeons 222 and 223. Nevertheless, results for all subjects showed evidence of hysteresis.

LINOP and the EDM were compared to provide a quantitative account of the present data. Log initial-link response ratios were computed for every block of six cycles in each session (i.e., twelve blocks per session), then averaged across replications, giving a total of 384 data points (12 x 32) for each subject. The models were then fitted by obtaining parameter estimates that maximised the variance accounted for in the data. For LINOP the parameters included Δ_S, which determined the rate of learning across sessions, and Δ_R, which determined the rate of learning within sessions. There was also a sensitivity exponent, q, in the function determining the asymptotic value of a delayed reinforcer, $V = 1 / (c + d^q)$, where delay is d seconds and c is an additive constant which was set equal to 0 for the fits presented here (see Grace, 2002a, Equation 4). An additive bias parameter, log b, was also used.

Parameter estimates that maximised the variance accounted for were obtained through nonlinear optimisation (Microsoft Excel© Solver).

Figure 17 shows the obtained log initial-link response ratios as a function of LINOP predictions for the individual block data (session 12ths). Slopes for best-fitting regressions are also shown, and the slopes are close to 1.0, suggesting that the LINOP model captured the overall trends in the data. Averaged across subjects, the LINOP model accounted for 91% of the variance in the session 12h-data. However, there is some evidence of sigmoidal curvature for some subjects in Figure 17 indicating that the LINOP predictions deviate systematically from the obtained data. For example, pigeons 221, 224, 192 and 194 appear to have obtained data that follow a trend that begins below the regression line at low predicted values and as the predictions increase, falls above the regression line. These subjects also have the most pronounced bubble between series in the session data (Figure 16), and suggest that LINOP struggles to describe hysteresis effects in response allocations. Parameter values for the fits of the LINOP model to the session-12th data are listed in Table 4.

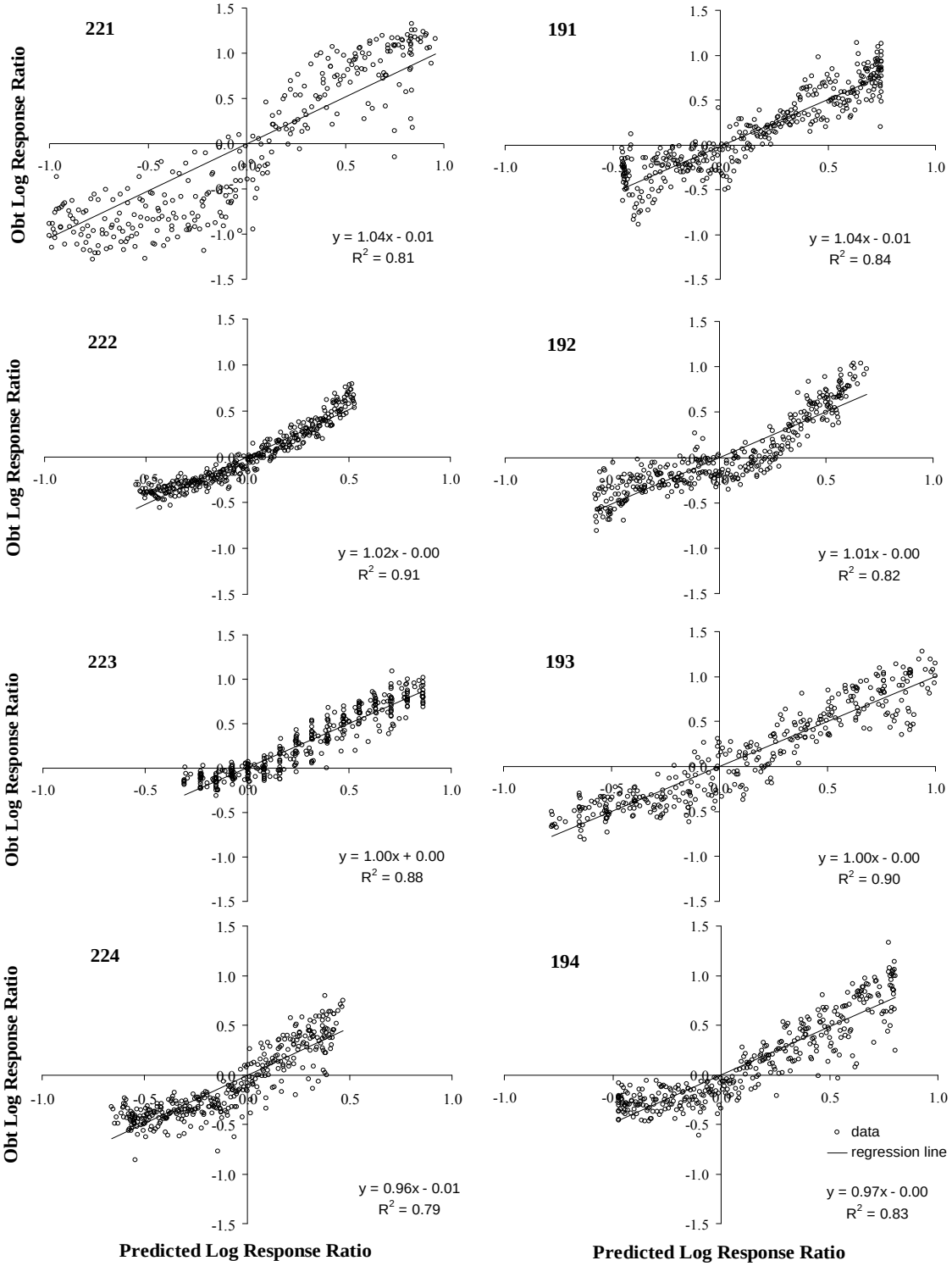


Figure 17. The session-12th obtained log initial-link response ratios as a function of LINOP-predicted log immediacy ratios for both ascending and descending series, averaged across replications for individual subjects. Included are the regression lines, associated best fitting r^2 values and linear regression parameters.

Table 4.

LINOP Parameter and VAC Values for Fits to Session 12th Data.

Pigeon	Session VAC	q	log b	Δ	Δ s	Session 12th R^2
221	0.89	2.25	-0.19	0.13	0.54	0.81
222	0.95	0.91	-0.01	0.17	1.00	0.91
223	0.96	1.04	0.32	0.58	0.27	0.88
224	0.89	1.09	-0.05	0.13	0.36	0.79
191	0.87	1.41	0.23	0.03	1.00	0.84
192	0.89	1.18	0.04	0.17	0.38	0.82
193	0.95	1.7	0.25	0.37	0.28	0.9
194	0.95	1.23	0.18	0.05	1.00	0.83
Average	0.92					0.85

The dashed lines in Figure 16 show the whole session average values (obtained and predicted by the LINOP model) as a function of the log immediacy ratio for both the ascending and descending series. Predicted values were calculated by averaging across the predicted values for the session-12th data. Overall, LINOP provided a reasonably good account of the data, accounting for 85% of the variance. LINOP was able to predict the separation between ascending and descending series in the full-session data, corresponding to the hysteresis effect. Additionally, LINOP was able to capture some of the nonlinearity in the full-session data (see pigeons 221, 224 and 194). However, LINOP appears to fail in capturing some of the patterns of hysteresis. In particular, the response patterns of pigeons 221, 224, 191, 193 and 194 appear to show evidence of little change in response ratios in the beginning of the ascending series. LINOP seems only able to capture this effect for pigeons 191 and 194 and fails to describe the hysteresis in the ascending series for pigeons 221, 224 and 193.

Figure 18 shows the extended decision model fitted to the session 12th data. The criterion value was calculated for each session as the average of the log programmed intervals between stimulus transitions (i.e., initial-link onset to terminal-link entry, and terminal-link entry to reinforcement). The probability of the relative current delay being judged "short" relative to the criterion was then used in the prediction of preference for the session-12th data. The maximum and minimum response strengths for both alternatives were initially set equal to 1.0 and 0.01, respectively. Solver was used to obtain best-fitting values of the standard deviation (σ), learning rate parameter for the terminal links (Δ) and between-session changes (Δ_S), as well as an additive bias parameter (log b). The parameters Max and Min were fixed at 1.0 and 0.01, respectively, whilst β was set at 1.0 and Log C held at the geometric mean of the delays 0.903. Thus, both the EDM and LINOP had four free parameters. Parameter values for the EDM fits to the individual data are listed in Table 5.

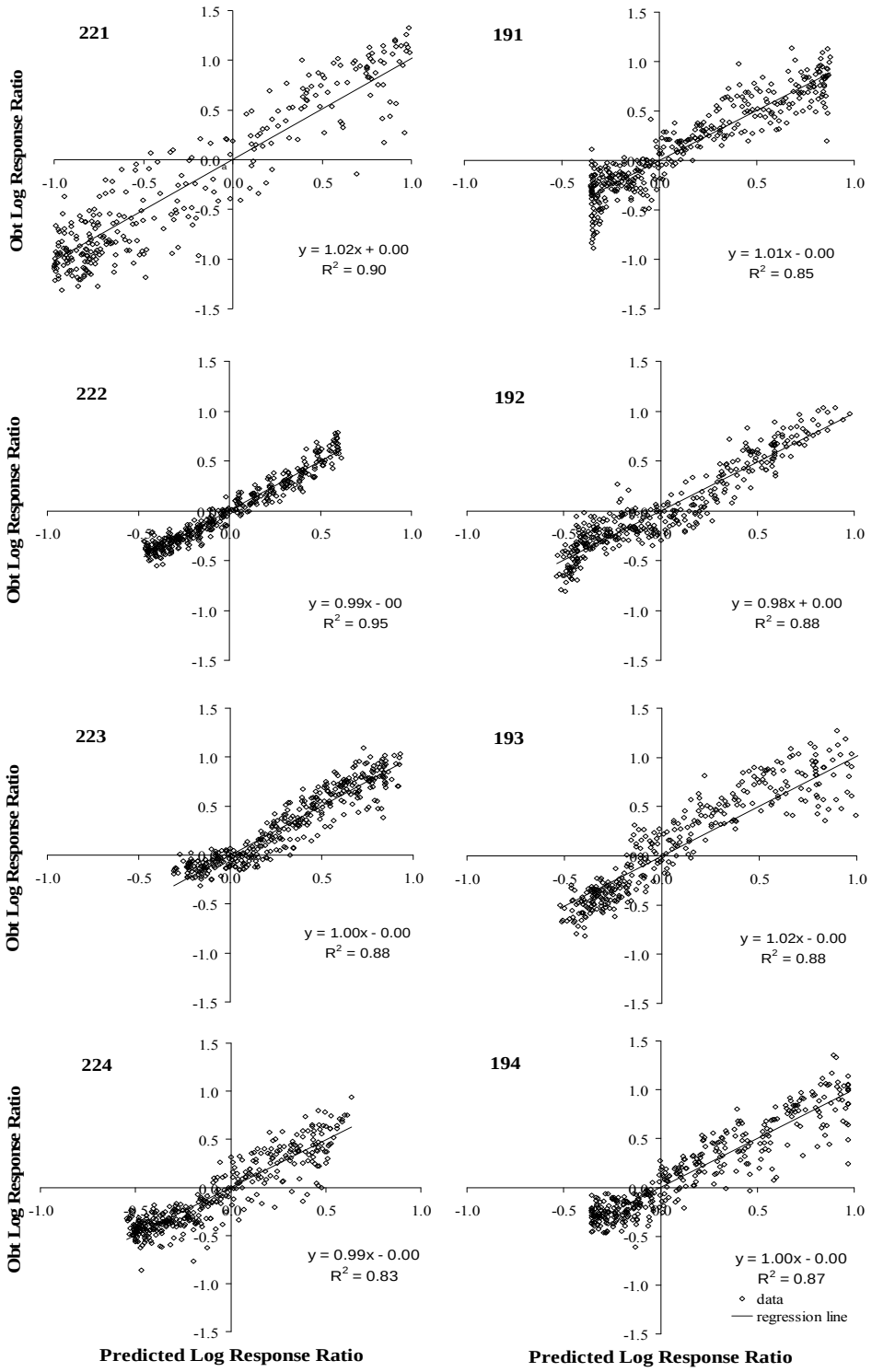


Figure 18. The session-12th obtained log initial-link response ratios as a function of EDM predicted log immediacy ratios for both ascending and descending series, averaged across replications for individual subjects. Included are the regression lines, associated best fitting r^2 values and linear regression parameters.

Table 5.

EDM Parameter and VAC Values for Fits to Session 12th Data.

Pigeon	Session VAC	Log *C*	σ	Log *b*	Δ	Δ s	Session 12th R^2
221	0.97	0.90	0.09	-0.64	0.21	0.41	0.90
222	0.99	0.90	0.34	-0.06	0.10	1.49	0.95
223	0.97	0.90	0.34	0.31	0.27	0.45	0.88
224	0.93	0.90	0.26	-0.16	0.11	0.36	0.83
191	0.89	0.90	0.17	0.04	0.03	0.69	0.85
192	0.95	0.90	0.22	-0.09	0.18	0.26	0.88
193	0.92	0.90	0.16	0.04	0.26	0.22	0.88
194	0.96	0.90	0.25	0.06	0.04	1.50	0.87
Average	0.96						0.88

Figure 18 shows obtained log initial-link response ratios as a function of EDM predictions for the session-12th data. The best-fitting regression lines are also shown. Overall, the EDM did a good job of describing the session-12th data, accounting for an average of 88% of the variance. The regression slopes were also all close to 1.0. However, there is some evidence of curvature in the scatterplots that indicate that predictions of the EDM, like those for LINOP, sometimes deviate systematically from the obtained values. For example, pigeons 192, 193 and 223 appear to have obtained data that follow a trend that begins below the regression line at low predicted values and as the predictions increase fall above the regression line.

The solid lines in Figure 16 show the resulting session average values as a function of the log immediacy ratio. The EDM provided an excellent account of the data, with an average VAC of 96% for the full-session data. The EDM provided a good description of results for subjects in which there was a clear separation between

the ascending and descending series, as well as when the series nearly superposed. For example, pigeon 224 has a distinct separation between series, while pigeon 223 has almost identical curves for the ascending and descending series. Compared to the LINOP predictions, Figure 18 appears to show the EDM is able to capture both patterns of responding. Moreover, the EDM also appears to capture hysteresis in both ascending and descending series. This is most evident in subject 221, where the obtained and predicted curves become flatter at the start of both ascending and descending series. In addition, the EDM seems to be a good approximation of more linear patterns of response allocation, for example pigeons 222 and 223.

Comparing the model fits, those for the EDM were overall superior, with higher VAC for 7 of 8 pigeons for the full-session data, and for 6 of 8 (with one tie) for the session 12th data. As the models have the same number of free parameters, this suggests that the EDM may provide a better description of response allocation for the present data.

However, even if two models have the same number of parameters, one may have greater flexibility in terms of being able to predict a greater range of outcomes (Pitt, Myung, & Zhang, 2002). If so, the model may account for a higher percentage of variance than a competitor because of this flexibility. Thus to determine whether the EDM and LINOP differed in terms of flexibility, both models were fit to simulated data generated by each model. The simulated data were obtained by adding random noise (distributed uniformly between -0.1 and 0.1) to the predicted values when each model was fitted to the average session-12th data. If either model is more flexible, then it should provide not only the best account of simulated data generated by that model, but an equal or better account of data generated by the other model as well. For simulated data generated from the EDM, the EDM and LINOP accounted

for 93.5% and 86.4% of the variance, respectively. For simulated data generated from LINOP, the EDM and LINOP accounted for 87.1% and 93.5% of the variance. In both cases, the model that generated the simulated data provided the better fit. This suggests that there is no difference in flexibility between EDM and LINOP. Therefore, it appears that the EDM provides a better overall account of the present data.

Evidence for the terminal-link effect – a stronger preference for the shorter terminal-link delay when the absolute values of the delays increase with their ratio held constant - was investigated. Figure 19 shows the obtained log response ratios (full session) as a function of the absolute value of the log immediacy ratio for individual subjects, separately, according to whether the terminal-link FI schedule for the right alternative was less than or greater than 8 s. Each data point represents an average across the ascending and descending series. As the 8-s duration was the midpoint of both series, the log immediacy ratios formed pairs with equal absolute values. The terminal-link effect predicts that sensitivity to the log immediacy ratio, as measured by the slope of the generalized-matching function of the log immediacy and log response allocation, should be greater when the right terminal-link schedule was greater than 8 s compared to when it was less than 8 s. Figure 19 shows that for all subjects the >8 s log response ratios had a greater slope than the corresponding < 8 s response ratios. Thus the data exemplified the terminal-link effect, that is, preference was more extreme with longer absolute terminal-link duration as relative duration was held constant. Figure 20 shows the group data and the model predictions.

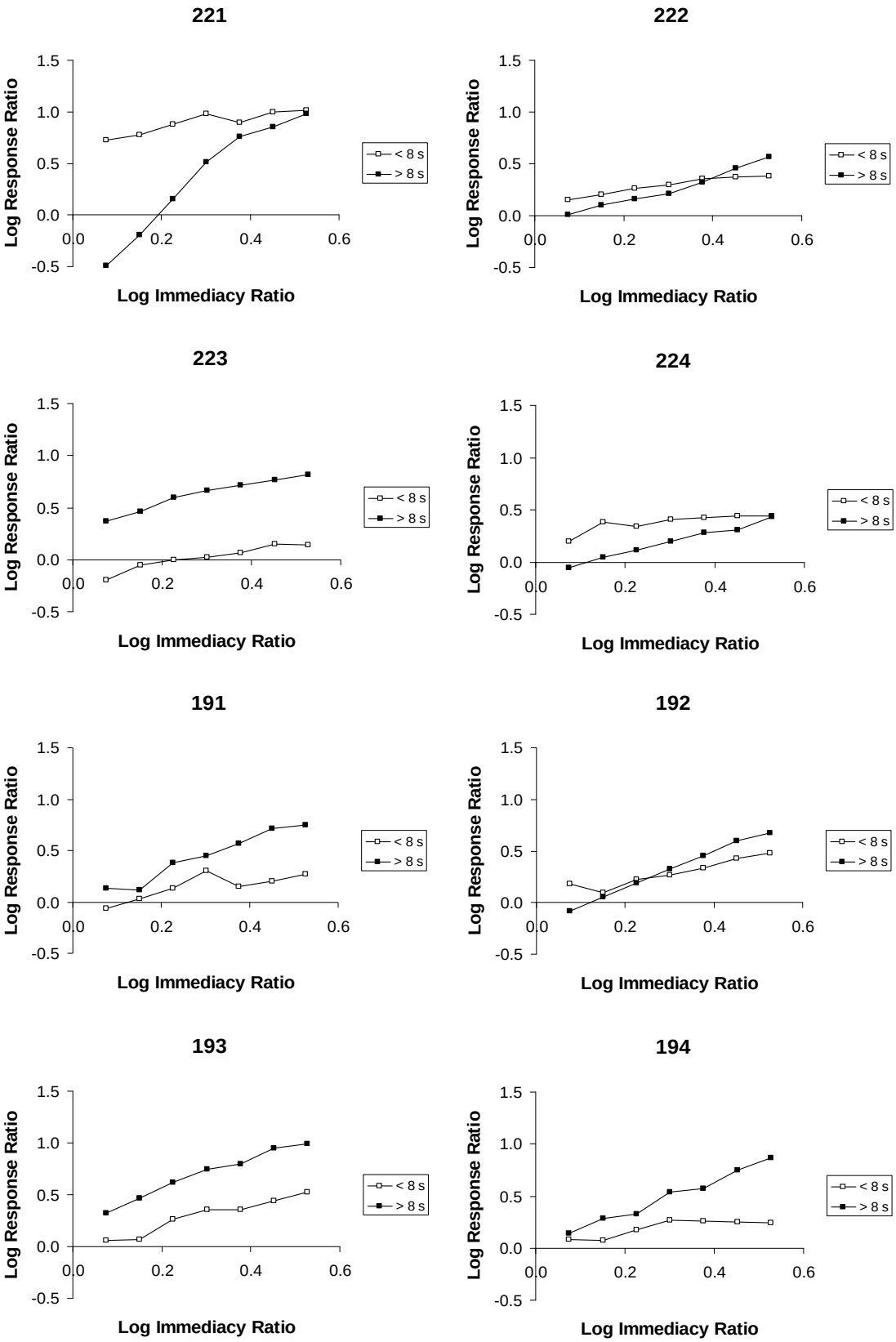


Figure 19. Log initial-link response ratios as a function of log terminal-link immediacy ratios for which the right terminal-link FI schedule was greater than or less than 8s, for both ascending and descending series, averaged across replications for individual subjects.

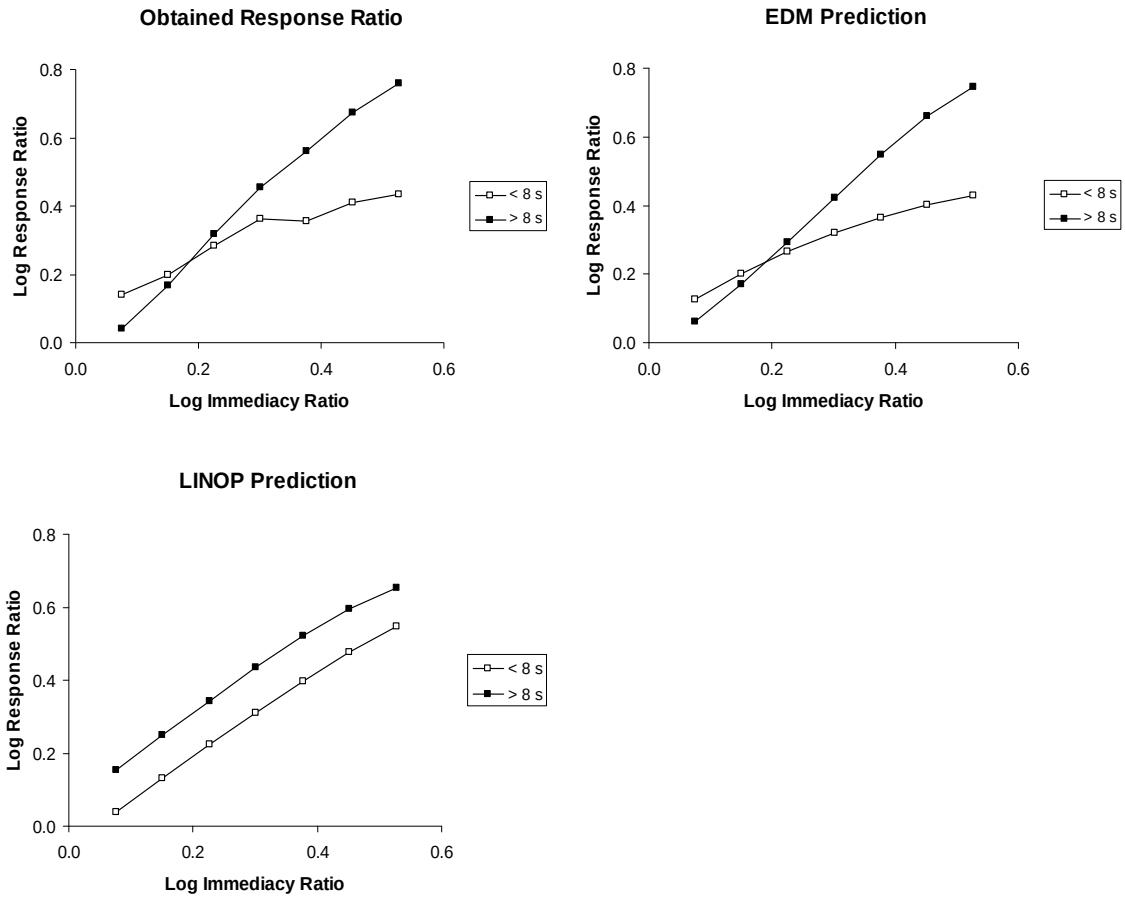


Figure 20. Obtained log initial-link response ratios (top panel) and predictions of EDM (lower left panel) and LINOP (upper right panel) as a function of log terminal-link immediacy ratios for which the right terminal-link FI schedule was greater than or less than 8s, for both ascending and descending series, averaged across replications and subjects.

Figure 20 shows the group averages for predicted and obtained log response ratios for both series when the terminal-link duration was less than or greater than 8 seconds (upper left panel). Like the individual data, the obtained data show steeper slopes in the >8s than the <8s log response ratios. The EDM predictions in the lower left panel (obtained by averaging across predictions for the ascending and descending series in Figure 18) show the same pattern as the obtained data. However, the

corresponding LINOP predictions have parallel slopes for the two sets of conditions, indicating that LINOP failed to predict the terminal-link effect.

Discussion

The present study explored how initial-link response allocation in concurrent chains changed when one terminal-link delay followed an ascending and descending sequence across sessions while the other remained constant. The goal was to test predictions of two models for acquisition in concurrent chains: the extended decision model and the LINOP model (Grace, 2002a). The decision model had previously been applied only to situations in which the terminal links changed unpredictably across sessions. Here, a proportion of the change in response strength within a session is assumed to be retained at the start of the next session.

The terminal-link schedule for the left alternative was always FI 8 s, while the right terminal-link schedule varied between FI 2 s and FI 32 s according to a geometric series. Two predictions of the EDM were evaluated: That a hysteresis or carryover effect would be obtained, and that response allocation would be more sensitive to changes in the immediacy ratio at higher absolute terminal-link durations (see Figure 14). Both predictions were confirmed.

For all subjects, scatterplots of the log initial-link response and log immediacy ratios showed a gap or bubble between data for the ascending and descending series (see Figure 16). This phenomenon occurred because the series tended to converge at the extreme immediacy ratios, whereas for intermediate ratios the log response ratio tended to favour the left initial link to a greater extent during the

descending series. As the descending series began after the right-key delay expected to produce maximal preference for the left key (32 s), the left-key bias during the descending series represents a hysteresis effect. Both LINOP and the EDM predicted the hysteresis effect.

This result is similar to that reported by Field, Tonneau, Ahearn and Hineline (1996), who studied pigeons' choices between fixed ratio (FR) 30 and variable ratio (VR) 60 terminal links in concurrent chains. Across successive phases of their experiment, the VR distribution was manipulated such that the minimum response requirement was changed according to an ascending and descending series. Preference for the VR alternative tracked the minimum requirement; a requirement of 1 produced a strong preference for the VR terminal link, and this preference decreased as the requirement was increased up to 15. Field et al. found that for a given minimum requirement, preference for the VR alternative was greater on the ascending than descending series, which is analogous to the hysteresis effect reported here. However one difference is that each phase in Field et al.'s experiment lasted for 11 sessions. Thus, despite the differences in procedure (interval vs ratio schedules; schedules changed after 1 and 11 sessions), both experiments produced similar hysteresis effects. It is unknown whether such hysteresis depends on how frequently the terminal links are changed.

Overall, the EDM provided a very good account of the data in quantitative terms, accounting for an average of 88% of the variance in the session-12th data and 95% of the variance in the session data. These are somewhat higher than the corresponding values for LINOP (85% and 92%), as well as for the fits of the original version of the decision model to Grace and McLean's (2006) data (73% and 84%). However, it is worth noting that there was some evidence of systematic deviation in

the obtained vs predicted scatterplots for the session-12th data (see Figure 18), indicating that the EDM was unable to capture all of the trends in the data.

Preference was also examined for the shorter terminal link when overall terminal-link duration increased with the immediacy ratio held constant. This result is known as the terminal-link effect, and has been one of the most widely-studied phenomena in concurrent chains, having been obtained when terminal links differ in terms of reinforcer magnitude (Navarick & Fantino, 1976) and probability (Spetch & Dunn, 1987), as well as immediacy (Grace, 2004; Grace & Bragason, 2004; MacEwen, 1972; Williams & Fantino, 1978). In the present experiment, the delays were geometrically spaced so the ratios between 1:1 and 4:1 were the reverse of those between 1:4 and 1:1. Thus sessions were compared in which the delays were either both less than 8 s, or both greater than 8 s, with the ratio of delays held constant. For all subjects, the slope relating log response allocation to the log immediacy ratio was steeper when the delays were greater than 8 s (Figure 19). This difference between slopes is congruent with the terminal-link effect, where response rates are more extreme for increasing terminal-link delays. This result was predicted by the EDM, but not LINOP (Figure 20) and also supports the addition of the initial-link delay to the calculation of the criterion. Thus, these results show that the extended decision model can be applied effectively to a situation in which terminal-link delays change systematically across sessions.

Experiment 5

The goal of Experiment 5 was to 1) test whether the terminal-link effect (that is, when absolute delay increases with relative delay held constant, response allocation for the richer alternative becomes more extreme) would be obtained in a PRBS procedure, and 2) examine whether the criterion in the EDM increases when terminal-link duration is increased. In this experiment, the initial links were constant and the terminal-link schedules were changed across sessions in a quasi-random binary sequence. There were two types of conditions, in which the terminal links were overall short (8 s vs 4 s or 16 s) or long (16 s vs 8 s or 32 s). To test whether the EDM could account for the terminal-link effect, the criterion was estimated separately for the long and short conditions. It is predicted that current session sensitivities from multiple regressions of log response allocation and log *C* values would be greater in the long condition than in the short condition.

Method

Subjects

Four pigeons of mixed breed, numbered 181, 182, 183 and 184, served as subjects and were maintained at 85% of their free-feeding weight plus or minus 15 g through appropriate post-session feeding. Pigeons were housed individually in a vivarium with a 12h:12h light/dark cycle (lights on at 0600), with water and grit

freely available in the home cages. All pigeons were experienced with rapid-acquisition concurrent-chains procedures.

Apparatus

Four standard three-key operant chambers, 32 cm deep x 34 cm wide x 34 cm high, were used. The keys were arranged in a row 21 cm above the floor. In each chamber there was a houselight located above the centre key that provided general illumination (except during reinforcer delivery), and a grain magazine with an aperture centred 6 cm above the floor. The magazine was illuminated when wheat was made available. A force of approximately 0.15 N was necessary to operate each key. Each chamber was enclosed in a sound-attenuating box, and an attached fan provided ventilation and white noise. Experimental events were controlled and data recorded through a microcomputer and MEDPC® interface located in an adjacent room.

Procedure

For all pigeons, training started immediately in a concurrent-chains procedure. With few exceptions, sessions occurred daily and at approximately the same time (1000 hours). Sessions ended after 72 initial- and terminal-link cycles had been completed or 70 min, whichever occurred first. At the start of a cycle, the side keys were illuminated white to signal the initial links. An entry was assigned pseudo-randomly to the left or right terminal link with the constraint that in every 6 cycles, 3 entries occurred to each terminal link. An initial-link response produced an entry into a terminal link provided that: (a) it was made to the preselected key; (b) an interval

selected from a VI 10-s schedule had timed out; and (c) a 1-s changeover delay (COD) was satisfied – at least 1 s had elapsed following a response to a lighted key to the side for which terminal-link entry was arranged.

The VI 10-s initial-link schedule did not begin timing until the first response had occurred in each cycle, to allow any pausing after the completion of the previous terminal link to be excluded from initial-link time. The VI 10-s schedule contained 12 intervals constructed from an exponential progression (Fleshler & Hoffman, 1962). Separate lists of intervals were maintained for cycles in which the left or right terminal link had been selected, and were sampled without replacement so that all 12 intervals would be used three times for both the left and right terminal links each session.

When a terminal link was entered, the colour of the side key was changed (left key to red, right key to green) while the other key was darkened. Terminal-link responses were reinforced according to FI schedules. When a response was reinforced all lights in the chamber were extinguished, and the grain magazine raised and illuminated for 3 s. The centre key was unused.

The experiment consisted of two conditions. In the short condition, the left terminal link was always 8s and the right terminal link was either 16 or 4 s, while in the long condition the left was always 16s and right terminal link was either 32 or 8 s. For both conditions, terminal-link delays changed each session in a pseudo-random binary sequence. The design of the experiment was ABA, with pigeons 181 and 182 experiencing a short-long-short order of conditions, while pigeons 183 and 184 experienced a long-short-long order. Pigeons experienced 62, 31 and 31 sessions in the three conditions, respectively, except for pigeon 182 that received only 27 sessions in the second condition due to an equipment failure.

Results

To assess the relationship between response allocation and the immediacy ratios in the current and prior sessions, I used a generalized-matching model:

$$\log\frac{B_{0L}}{B_{0R}} = a_0 \log\frac{1/D_{0L}}{1/D_{0R}} + a_1 \log\frac{1/D_{1L}}{1/D_{1R}} + a_2 \log\frac{1/D_{2L}}{1/D_{2R}} + \ldots + \log b \qquad (29)$$

where B refers to initial-link response rate, subscripted for choice alternative (L and R), and D represents terminal-link delay. The parameters a_0 through a_3 quantify the sensitivity to reinforcer immediacy (the reciprocal of delay) for the current session (lag 0) and the preceding three sessions (lag 1-3), and log b is a bias parameter. A series of multiple regressions were conducted in which Equation 29 was applied to the data for individual pigeons from the last 20 sessions of each condition because subjects experienced different numbers of sessions and to maintain consistency with earlier experiments. The lag coefficients are shown in Table 6.

Table 6.

Multiple Regression Lag Coefficients

181				182			
	Short	Long	Short		Short	Long	Short
Lag0	**0.56**	**1.89**	**0.82**	Lag0	**0.49**	**1.37**	**1.22**
Lag1	**0.15**	0.06	-0.01	Lag1	0.15	0.01	0.01
Lag2	0.04	0.14	-0.06	Lag2	0.02	0.04	-0.02
Lag3	0.01	0.04	-0.06	Lag3	0.14	0.25	**0.24**
R^2	0.86	0.94	0.86	R^2	0.68	0.9	0.94

183				184			
	Long	Short	Long		Long	Short	Long
Lag0	**2.40**	**1.06**	**2.20**	Lag0	**1.05**	**1.01**	**1.67**
Lag1	0.27	-0.06	0.16	Lag1	0.14	0.17	0.06
Lag2	0.12	0.06	0.04	Lag2	0.15	0.01	0.15
Lag3	-0.04	-0.03	-0.03	Lag3	0.12	0.2	-0.01
R^2	0.94	0.62	0.97	R^2	0.87	0.82	0.96

Bold $p<0.05$

For each subject and condition, response allocation was sensitive to the immediacy ratio in the current session using a multiple regression where Lag0 to lag3 was regressed on the current session's log response allocation, as indicated by *a* Lag 0 coefficient that was positive and statistically significant. Higher lag coefficients (lags 1-3) were generally not significant. There were two exceptions: the Lag 1 coefficient was significant ($a_1 = 0.15$) for pigeon 181 during the initial exposure to the short-duration condition, and the Lag 3 coefficient ($a_3 = 0.24$) was significant for pigeon 182 in the replication of the short-duration condition. In all other cases, coefficients for Lags 1 through 3 were not significant. This shows that response allocation was sensitive to the current-session reinforcer contingencies, similar to previous studies with the PRBS procedure (Grace et al., 2003).

Figure 21 shows a comparison of current session (Lag 0) sensitivity between the short and long conditions. For each pigeon and replication, Lag 0 sensitivity was greater for the long than short condition. Averaged across subjects and replications, Lag 0 sensitivity was significantly greater in the long condition (M=1.73) than the short condition (M=0.90), $t(3)=3.59$, $p<0.05$. This result is congruent with the "terminal-link effect" – an increased preference for the preferred alternative (the shorter terminal link delay) becomes more extreme as delays increase with the ratio held constant (McEwen, 1972) – which is obtained when terminal-link schedules change unpredictably across sessions.

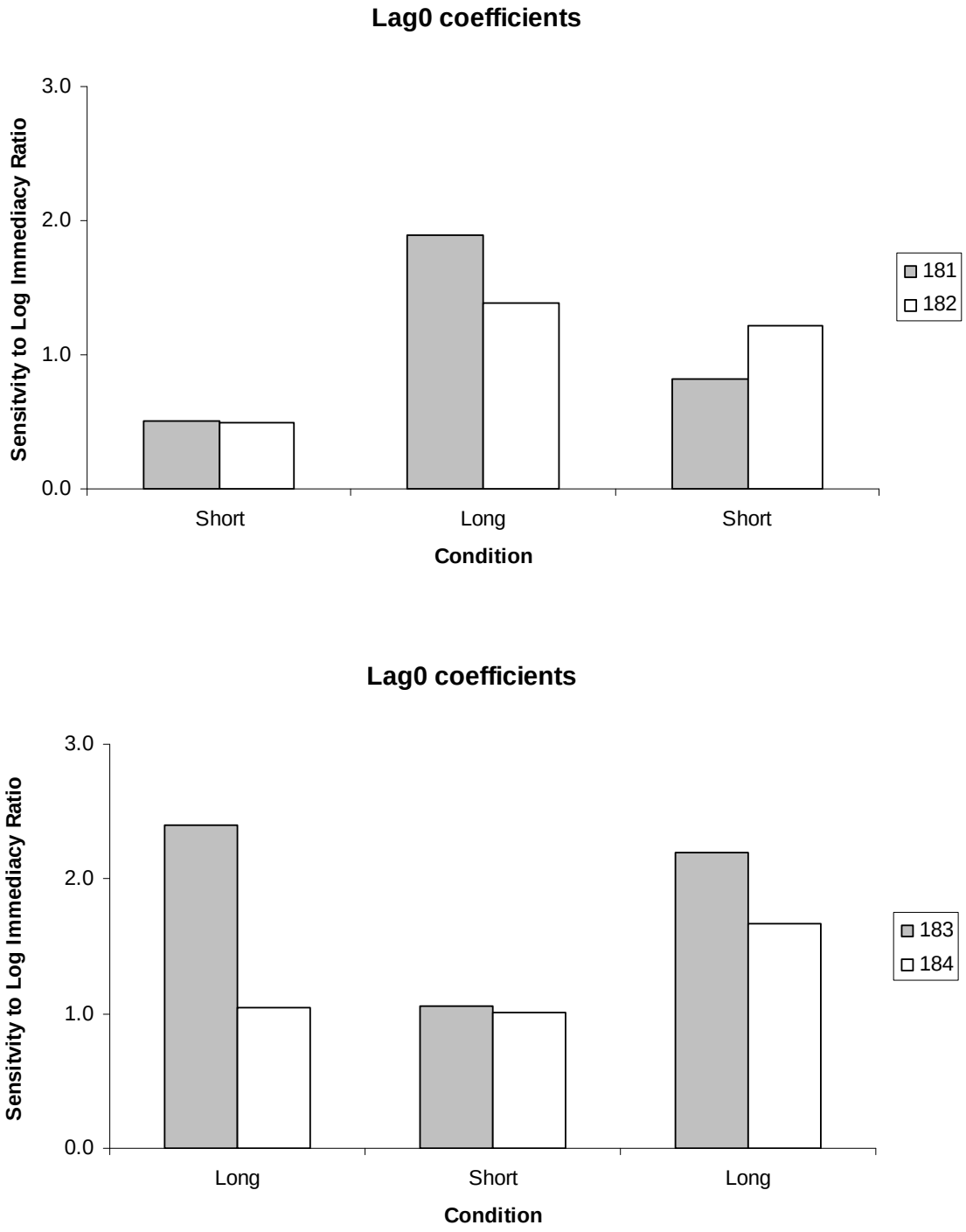


Figure 21. Sensitivity to the current-session (lag 0) terminal-link delay ratio for individual subjects in each condition.

Next, the EDM was fitted to the session-12th data for each pigeon's data for the last two conditions (12 data points in each session). For each pigeon, there were 240 data points from each condition (480 total). These data were fitted with σ, Δ as free parameters, and log b and Log C separately for each condition. The exponential weighting parameter β was fixed at 1.0.Thus, a total of six free parameters were fitted to 480 data points for each pigeon. The maximum and minimum response strength values were held constant at 1.0 and 0.01, respectively. The starting response strength was set at 0.1 (which is the geometric mean of the minimum and maximum values) for each alternative. In each case, values of parameters that maximised the variance accounted for in the session-12th data were obtained. Table 7 lists the resulting parameter values for each pigeon and variance accounted for by the EDM.

Table 7

Decision Model Parameter Estimates

Pigeon	VAC Full Session	VAC Block	Log *C* Short	Log *C* Long	σ	Log *b*	Δ
181	0.89	0.77	1.13	1.19	0.05	-0.41	0.06
182	0.89	0.75	0.92	1.07	0.18	-0.23	0.14
183	0.96	0.90	0.97	1.09	0.19	-0.21	0.36
184	0.88	0.71	0.85	0.84	0.33	0.28	0.44
Average	0.91	0.78	0.97	1.05	0.19	-0.14	0.25

Table 7 shows that the decision model provided a good account of the data. Averaged across subjects, the model accounted for 78% and 91% of the variance in the block and full-session data, respectively. We calculated the sensitivity to the current-session delay ratio predicted by the decision model by regressing full-session predictions on the log delay ratios. For all subjects, predicted sensitivities were greater for the long condition compared to the short condition, *M*'s = 1.25 and 1.68, confirming that the decision model accounted for the terminal-link effect. These were close to the corresponding obtained values, which averaged 1.12 and 1.62, respectively.

Estimates of Log *C* are also shown in Table 7. Estimates were greater for the long condition compared to the short condition for 3 of 4 individual pigeons (for the fourth, values were nearly equal). Averaged across pigeons, Log *C* values were 1.05 and 0.97 for the long and short conditions, respectively. An increase by a factor of 2 in the criterion implies a difference of Log (2) = 0.301 in the Log *C* values. However, the obtained difference was only 0.08, which translates to an increase by a factor of 1.20. Thus, estimates of the criterion increased in the long condition by less than a factor of 2, consistent with predictions of the decision model. The EDM fits to the session 12th predictions are illustrated in Figure 22.

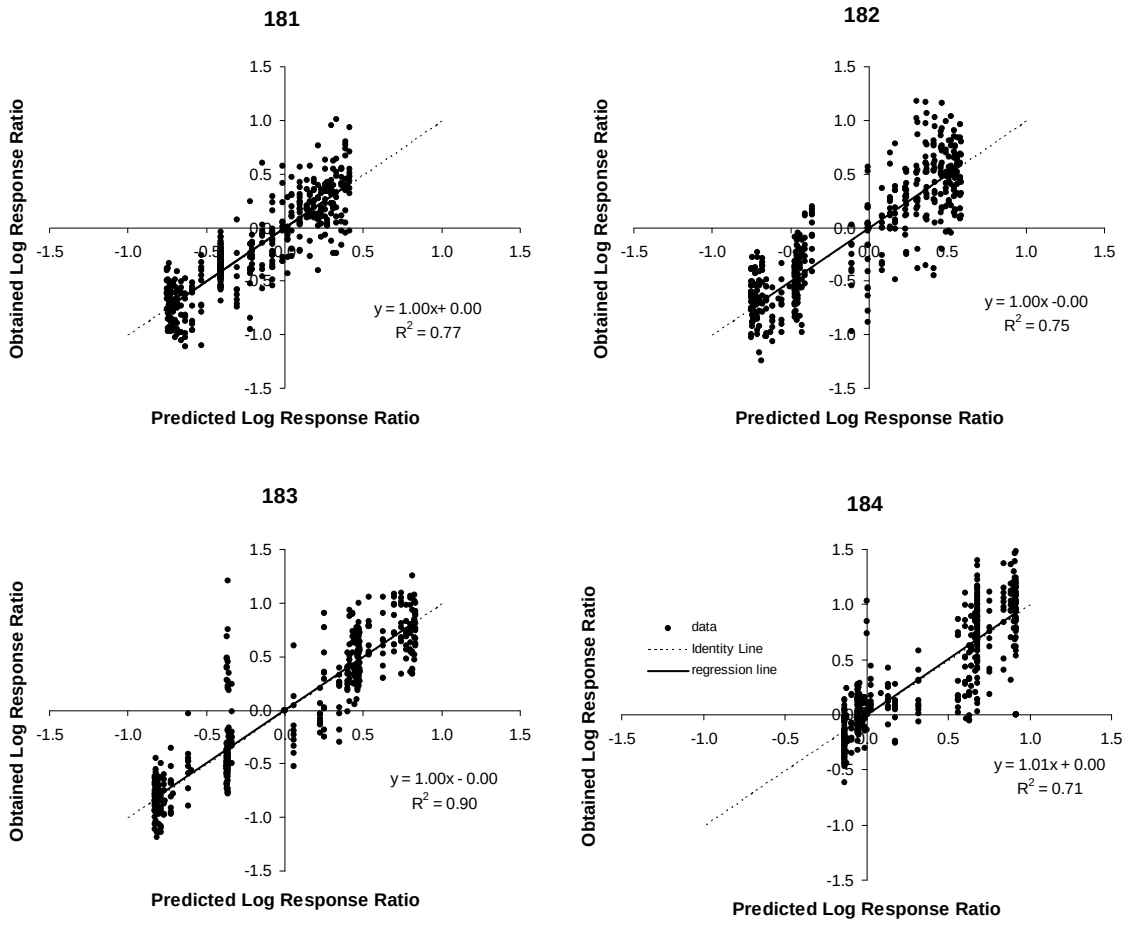


Figure 22. Scatterplots of obtained and predicted log response ratios for session 12ths data. Also shown are the regression and identity lines.

Scatterplots of obtained versus predicted session-block data are shown for individual pigeons in Figure 22. For all pigeons the best fitting regression of log response allocation falls on or very close to the line of identity. This is confirmed by an average slope of 1.0 and the average bias 0.0 suggesting that the obtained data did not deviate systematically from the predicted values. For all pigeons, the slope of the best fitting linear regression line falls on the line of identity for the session-12th plots. Visual inspection appears to show equal spread on both sides across the regression

line indicating the EDM is describing the overall trend for response allocation within sessions.

Discussion

This experiment examined response allocation in concurrent chains when the initial-link schedule was constant and the terminal-link delays changed across sessions according to a pseudorandom binary sequence. Overall terminal-link duration was manipulated across conditions: In the short condition the left terminal link was FI 8 s and the right terminal link was either FI 4 s or FI 16 s, whereas in the long condition the left terminal link was FI 16 s and the right terminal link was either FI 8 s or FI 32 s. Multiple regression analyses found that response allocation was sensitive to the delays in the current session, with higher sensitivity when the delays were long than when they were short. This replicates the terminal-link effect (MacEwen, 1972; Williams & Fantino, 1978) in a PRBS design. The decision model (Grace & McLean, 2006) provided a good description of the data, accounting for an average of 78% and 91% in the block and full-session data, respectively, and was able to predict the terminal-link effect for all subjects. Notably, estimates of the criterion (Log *C*) increased in the long condition, but by less than a factor of 2, and were overall close to the expected values based on the programmed schedule durations. These results confirm that the terminal-link effect is a robust phenomenon and can be obtained in a rapid acquisition design, and provide additional support for the decision model as an account of response allocation in concurrent chains.

The terminal-link effect has long been considered one of the more intriguing phenomena in research on behavioural choice because it represents a clear violation of

Weber's Law, which typically characterizes performance in temporal discrimination paradigms (Gibbon, 1977). According to the decision model, the explanation for this violation is that the initial links also contribute to the criterion; if the criterion increased proportionally with terminal-link duration, then the model would predict invariant preference, consistent with Weber's Law.

The most influential account of the terminal-link effect has been Fantino's (1969) delay-reduction theory (DRT), which assumes that the conditioned reinforcement value of the terminal-link stimuli depends on the reduction in delay to reinforcement signalled by the terminal links, relative to the overall average delay from the onset of the initial links (*T*). Note that *T* is functionally similar to the criterion in the decision model, because it represents a comparator used to determine the stimulus value (DRT) or response strength for the initial link (decision model). To apply DRT to the present experiment, response allocation was calculated (as a percentage) for FI 4 s FI 8 s (short) and FI 8 FI 16 (long). For the short condition, *T* is computed as 10 s + (4 s + 8 s)/2 = 16 s. Thus the delay reductions associated with the FI 4 s and FI 8 s terminal links are 12 s and 8 s, respectively, giving a predicted response allocation of 0.60 (12 s / [12 s + 8 s]). For the long condition, *T* is 10 s + (8 s + 16 s)/2 = 22 s, and so the delay reductions for FI 8 s and FI 16 s are 14 s and 6 s, respectively, giving a prediction of 0.70 (14 s / [14 s + 6 s]). Thus DRT predicts that response allocation is more extreme in the long condition, exemplifying the terminal-link effect.

Note that the reason DRT makes this prediction is because *T* does not increase by a factor of two between conditions. If it did, then *T* would be 32 s in the long condition and predicted response allocation would be 0.60, equal to the short condition (and consistent with Weber's Law). This means that the explanation for the

terminal-link effect is fundamentally similar for the decision model and DRT: In both cases, the comparator (Log *C* for the decision model, *T* for DRT) increases less than proportionally with terminal-link duration.

Thus, although different in terms of quantitative details, the decision model and DRT share some strong similarities. Both models assume that choice in concurrent chains is determined by a process in which the attractiveness of a choice alternative is a function of the terminal-link delay relative to a comparator, which is based on both initial- and terminal-link durations. However, the decision model has the advantage of specifying response allocation as the outcome of a dynamic process.

Chapter 4: Archival Analyses

The rationale of this chapter is to compare the steady-state version of the EDM with existing models of steady-state choice. Specifically, the EDM will be evaluated against the leading models of choice: Contextual Choice Model (CCM: Grace, 1994) and the Hyperbolic Value Addition model (HVA: Mazur, 2001).

The archival data sets chosen were the same as analysed by Grace (1994) and Mazur (2001). These studies were based on the following criteria: a) minimum 4 data points for each subject, b) time based terminal-link schedules, either FI or VI, c) equal terminal-link magnitudes and d) obtained preference measures with a minimum range of 12%. In addition to these criteria, we also omitted conditions with zero initial links, and repeated conditions. The exclusion of zero initial links conditions was necessary as zero conditions make calculation of the criterion in the decision model problematic because the EDM assumes a concurrent-chains design and therefore positive schedule values for each link. As all models (CCM, HVA, and the EDM) made identical predictions for Squires and Fantino (1971), this data set was omitted from the analyses (in this study, the terminal links were equal across conditions and the initial links were varied; because the models all assume that the effects of relative terminal-link entry rates are described according to the Generalized Matching Law, their predictions are identical). The data sets that were affected were Davison (1983) (25 data points in total excluded due to zero initial-links), Fantino and Davison (1983) (6 data points excluded due to zero initial-links), MacEwen (1972) (16 data points excluded due mix of FI and VI conditions; VI conditions were omitted), Squires and Fantino (1971) (whole set deleted), and Williams and Fantino (1978) (24 data points

excluded due to replications). In total, 25 conditions were dropped from the original analyses, leaving 298 conditions consisting of 1463 data points in total to form the basis of the archival analyses. The final data sets included a range of procedures, independent and non-independent initial links, changeover and non-change over delays, response independent and response dependent terminal links, FI FI and VI VI schedules, and cued and un-cued terminal links.

The CCM and EDM model fits were calculated using Microsoft Excel Solver to maximise the variance accounted for by estimating the best fitting free parameters to the obtained data for each condition and each subject. Attempts to replicate Mazur's (2001) procedure to fit HVA using his original program were problematic because the Quick Basic © compiler he used was no longer available. When a similar open source program (Free Basic ©) was used this was found to be unsatisfactory, grossly under fitting the data. However, when Microsoft Excel solver © was used the model fits were within 1% - 3% of the reported values that used the same number of data points (Mazur, 2001). The likely differences were because of the different method of parameter estimation used by Solver, leading to slightly higher variance accounted for than previously reported for HVA (the algorithm used by Mazur, 2001, was based on a grid search and did not consider as many potential parameter values as the Solver algorithm). For all model fits, response-allocation data were scaled as choice proportions, to be consistent with the analyses reported by Grace (1994) and Mazur (2001). Table 8 lists the model fits across archival studies.

Table 8.

Archival Model Fits

Experiment	EDM	CCM	HVA
Alsop and Davison (1988)	0.91	0.91	0.90
Chung and Hernnstein (1967)	0.89	0.91	0.91
Davison (1976)	0.97	0.98	0.96
Davison (1983)	0.70	0.82	0.85
Davison (1988)	0.94	0.92	0.90
Davison and Temple (1973)	0.91	0.92	0.93
Duncan and Fantino (1970)	1.00	0.98	0.95
Dunn and Fantino (1982)	0.94	1.00	0.93
Fantino (1969)	0.98	0.99	0.94
Fantino and Davison (1983)	0.93	0.92	0.91
Fantino and Royalty (1987)	0.81	0.81	0.79
Gentry and Marr (1980)	0.77	0.78	0.81
Killeen (1970)	0.99	0.99	0.98
MacEwen (1972)	0.99	0.94	0.94
Omino and Ito (1993)	0.92	0.95	0.92
Preston and Fantino (1991)	0.80	0.75	0.76
Wardlaw and Davison (1974)	0.93	0.94	0.93
Williams and Fantino (1978)	0.92	0.96	0.96
Average	0.91	0.91	0.90

Table 8 compares the percentage of variance accounted for by predictions using HVA, CCM and the EDM. The calculations are rounded to 2 decimal places. The overall percentage of variance accounted for by each model is extremely close to previous reports (Grace, 1994; Mazur, 2001). The fits for CCM were within 1% of the reported values, while the HVA model fits were within 2-3% of earlier values, despite using a spreadsheet application rather than the earlier procedure that had separate Quick Basic © parameter estimation programs for each subject (Mazur, 2001).

All models fitted the data extremely well, with the EDM slightly outperforming both CCM and HVA. In most cases the EDM was similar to the fits for

the other models except for the Davison (1983) data. Table 8 shows that the EDM accounted for 10% less variance than either CCM or HVA. Although this data set had some conditions excluded due to 0-s initial links, these amounted to only 5 conditions out of 61, still making this experiment one of the largest in this re-analysis. Most likely, the difference between model fits for this data set is due to the extra fitted parameters used by both HVA and CCM. The EDM has only three free parameters (log b, $a1$, and σ) at steady-state, while both CCM and HVA have four (b, $a1$, $a2$, and k for CCM); (b, $a1$, a_T, and k for HVA). Therefore, the large number of data points and the numerous changes to both alternatives for initial- and terminal-link durations in the Davison (1983) data set would give an advantage to more complex models like CCM and HVA.

To assess whether systematic residuals were present, residual scores were pooled from the individual data. Figure 23 shows the residual values (obtained minus predicted choice proportions) plotted as a function of the predicted values. Predicted values were corrected for bias prior to this analysis. The correction procedure was as follows: predicted response allocation was fitted to the obtained data where bias was included as a free parameter. Bias was then eliminated by subtracting it from the predicted response allocation in each condition.

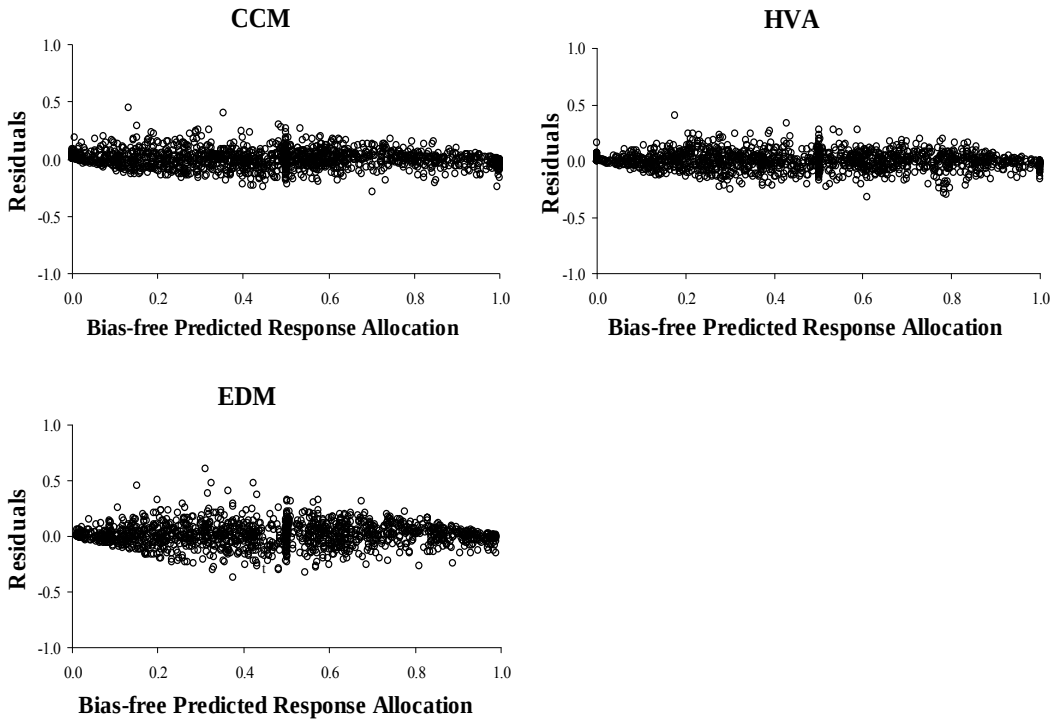


Figure 23. Residual components from 3 models as a function of bias-free log initial-link response ratios. Each unfilled circle represents a separate residual score from the archival analyses.

Figure 23 illustrates that CCM, HVA and EDM all have broad residual scatter across predictions. The predicted and residual scores were then entered into a polynomial regression in which the linear and cubic components were tested. Quadratic components were excluded because this function (U shape or inverted U shape) is not invariant under admissible transformations of the dependent variable (the L/R designation for log response ratios is arbitrary; see Sutton, Grace, McLean & Baum, 2008). Thus, only linear and cubic trends were analysed. To investigate whether the floor or ceiling effects were affecting residual components upper and lower limits were introduced and regressions were recalculated.

Table 9.

Linear and Cubic Residual Scores

Upper and Lower Limit Residual Scores

Models		Cutoffs		
		0.10	0.05	0.025
HVA				
	Linear	-0.02	-0.13	-0.01
	Cubic	0.01	-0.01	-0.01
CCM				
	Linear	0.02	0.00	0.01
	Cubic	0.00	-0.01	-0.01
EDM				
	Linear	0.02	0.04	0.05
	Cubic	0.01	-0.03	-0.04

Table 9 shows the beta weights for each model across the pooled data sets. The regressions were run using all of the available data, and with upper and lower limits. These were 0.1, 0.05 and 0.025 and were selected to test whether results were robust to potential ceiling and floor effects associated with percentage measures, by excluding the relatively extreme data points. Using a variety of upper and lower limits, no model had significant linear or cubic components in their residuals.

To examine whether there was any difference between schedule types experiments with VI VI and FI FI terminal-link schedules were separated, 12 with FI FI (Chung & Herrnstein 1967; Davison 1976, 1983, 1988; Davison and Temple 1973; Duncan & Fantino, 1970; Gentry & Marr, 1980; Killeen, 1970; MacEwen, 1972; Omino & Ito, 1993; Wardlaw & Davison, 1974; Williams & Fantino, 1978) and 6 within VI VI (Alsop & Davison, 1988; Fantino, 1969; Fantino & Davison, 1983;

Dunn & Fantino, 1982, Fantino & Royalty, 1987; Preston & Fantino, 1991). Specifically, terminal-link sensitivity, the *a2* parameter for CCM, a_T for HVA, and σ for EDM, was examined for invariance. Grace (1994) found CCM made different average estimates across schedule types, with much larger *a2* values in FI FI than VI VI schedules. This finding was replicated in the current analysis with higher average *a2* sensitivity values for the FI FI Schedules (2.37) than the VI VI schedules (0.77). However, the opposite result occurred for HVA. There was higher average a_T values in the VI VI schedules (1.23) than the FI FI schedules (0.67). This reversal also occurred for estimates of the EDM σ parameter, with higher average VI VI schedules (0.31) than FI FI schedules (0.28). The terminal-link sensitivity parameter estimates were compared using t tests to assess whether there were statistically significant differences between schedule types. For both CCM, $t(85) = -4.601$, $p<0.05$, and HVA, $t(85) = 3.05$, $p<0.05$, there was a significant difference between schedules, while there was no significant difference for the EDM, $t(85) = 0.687$, $p = 0.493$. Figure 24 plots the distribution of the equivalent terminal-link sensitivity parameters for each model.

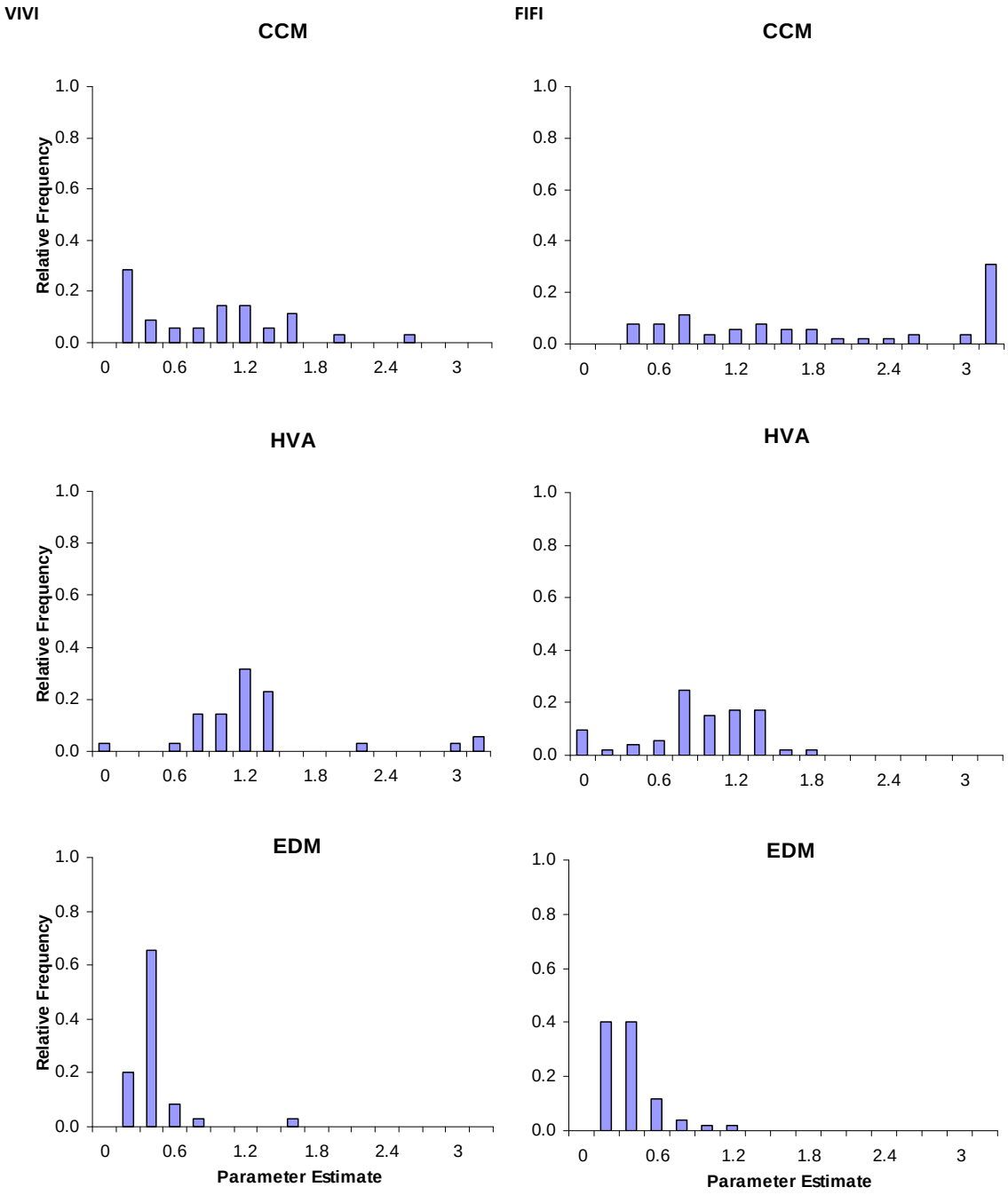


Figure 24. Histogram of relative frequency of terminal-link parameter estimates of archival analysis (Grace, 1994; Mazur, 2001) separating VI VI schedules and FI FI schedules.

Figure 24 shows the distribution for VI VI schedules (left panel) and FI FI schedules (right panel) for CCM, HVA and the EDM. Figure 24 illustrates that the distribution for both CCM and HVA were somewhat different for the VI VI and FI FI schedules. CCM appears to have greater spread of parameter estimates in FI FI schedules compared to VI VI, while HVA appears to have a more compact distribution in the FI FI schedules than VI VI. However, the EDM distribution appears to be more consistent than the other two models between schedule types, with similar positive skew and range of values for both schedules. Thus, it appears that the EDM parameter estimates of σ for changes in terminal-link delay to reinforcement in the archival data sets are relatively consistent across schedule types. Previous theoretical discussions have suggested that parameter invariance is an important property for quantitative models (Davison, 1987; Nevin, 1984), but no previous model has been able to demonstrate this property. Thus, the EDM appears to have unique and desirable qualities as a steady-state model for choice.

The results of the comparison between models show that EDM describes the data equally well as CCM and HVA in the archival analyses. Moreover, the parameter estimates for EDM appear to be constant across studies with different types of terminal links, indicating parameter invariance. Thus, the EDM appears to be a better account of the archival data than either of the previous models.

Chapter 5: General Discussion

The Temporal context of reinforcement has been typically seen as an influential factor on choice behavior (Fantino, 2001) which has traditionally challenged models for steady-state response allocation (The Matching Law, equation 1: Fantino, 1969, equations 6 and 7). In this thesis, the effect of changing the temporal context of reinforcement – overall initial- and terminal-link duration –was examined under conditions in which schedules changed each session in a concurrent chains procedure. The intention of these experiments was to examine predictions of an extended version of Grace and McLean's (2006) decision model (DM) in both transition and steady-state procedures. In particular, these models examined the initial- and terminal-link effects and the EDM was fitted to both situations. The steady-state version of the EDM was fitted to a set of archival data and was compared against two other models of steady-state response allocation.

Model Comparisons

Existing steady-state models have assumed different mechanisms for describing the effects of context in concurrent-chains procedures. For example, Delay Reduction Theory (DRT, equations 6 and 7) assumes that the relative reduction in expected time to reinforcement determines the effectiveness of terminal-link stimuli as conditioned reinforcers (Fantino, 1969), whilst the Hyperbolic Value Addition Model (Mazur, 2001, equations 9, 10, and 11) calculates a hyperbolic function of the delay between stimulus onset and reinforcement based on the value added from entering the terminal link. The Contextual Choice Model (CCM, equation 8), assumes

that the effectiveness of terminal-link reinforcement is modulated by the ratio of overall terminal- and initial-link duration, but the value of stimuli is determined independently of context, whereas the EDM (equation 21, 22, 23, and 24) assumes a combination of a linear-operator and categorical discrimination process based on the onset of stimuli correlated with reinforcement. Therefore, although there are similarities in the basic mechanisms between these models – the common assumption of using stimuli correlated with reinforcement – there are also differences in the construction of the calculations of response strength. To assess whether or not HVA, DRT and CCM could predict the bitonic initial-link effect, simulations similar to Figure 1 were conducted for each model.

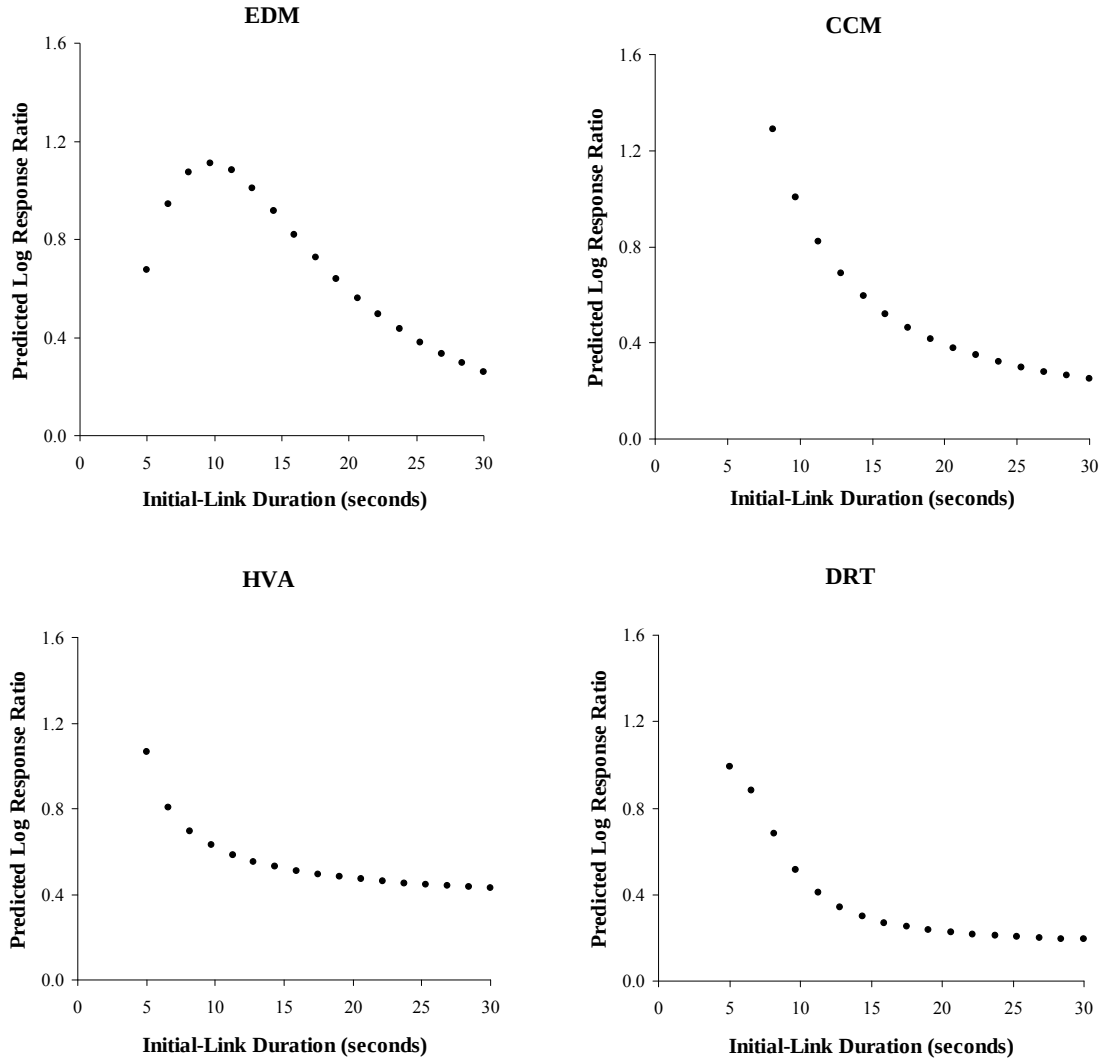


Figure 25. Initial-Link simulations for 4 models as the initial links were increased from 5s to 30s while the terminal-link durations were held constant at FI 10s and FI 20s.

Figure 25 shows the initial-link effect predictions for HVA, CCM, DRT and EDM. The parameters for each model are typical estimates from the data fits of the archival data. For CCM the parameter vales were; b = 1.2, $a1$ =0.5, $a2$ = 1.6, $a3$ = 0.8, k = 1.5, DRT; b =0.2, $a1$ = 1.0, $a2$ = 1.0, $a3$ = 1.0, k = 2.0, EDM; Log c = 1.0, σ =0.1, log b =0.0, Δ = 0.3, Max = 1.0, Min = 0.05, and HVA; b = 7.8, $a1$ = 0.4, a_T = 1.05, k =

1. The schedule type was assumed to be a variable-interval 12-step exponential distribution with initial-link durations ranging from 5s to 30s for both alternatives, and fixed interval terminal-link durations of 10s for the left and 20s for the right alternative. Figure 25 shows CCM, DRT and HVA have the expected decreasing monotonic function for log initial-link response ratios, as initial-link duration increases (Fantino, 1969). As the initial links are shortened, predicted response allocation continues to increase for these models. However, only the EDM predicts a bitonic function in both the ascending and descending series; response ratios increase then decrease as the initial-link durations increase. Experiment 1 confirmed this result. The prediction attributes the bitonic effect because of a reduction in the discrimination of short versus long delays when initial-link delays become very short.

Additional Model Comparison Issues

Although there have been numerous models of acquisition for concurrent-schedule performance, including Davison and Baum (2000), Mazur (1992), Myerson and Miezin (1980), Mazur (1996); the only acquisition model for concurrent chains, prior to the DM/EDM, was LINOP (Grace, 2002a). Results of Experiment 4 showed that although both LINOP and the EDM predicted hysteresis in response allocation when terminal-link schedules changed systematically according to an ascending and descending series, the EDM predicted the effects of overall terminal-link duration whereas LINOP did not.

The use of a Gaussian distribution, mean and a standard deviation to categorise delays into “short” or “long”, shares an obvious similarity with Signal Detection Theory (SDT). Yet there are important differences. The classical SDT

model assumes two Gaussian distributions, which correspond to the perceptual effects for trials in a detection task when the signal is present, and trials when signal plus noise is presented (Green & Swets, 1966). By contrast, the EDM assumes a single distribution corresponding to the history of initial- and terminal-link delays. In SDT, the criterion is the value of the likelihood ratio (i.e., the relative likelihood that the percept on a given trial was generated by the signal distribution) that must be exceeded. If the likelihood ratio is greater than the criterion (e.g., 1.0), the subject responds 'signal', otherwise the subject responds 'noise'. In the EDM, the criterion is the delay for which the subject is equally likely to classify as 'short' or 'long'. The outcome of a decision in SDT is a response, whereas for EDM the outcome is a change in the relative propensity to respond to the choice alternatives. Both SDT and EDM assume that the variability of the distributions is an important parameter influencing the accuracy of decisions. SDT calculates the discriminability between the two distributions based on separation between the distributions and their internal variability, which unlike the EDM, does not directly depend on the criterion, but calculates the actual hit rate and the false alarms to determine decision-making accuracy. Therefore, although the EDM shares some features of SDT, such as assuming a Gaussian distribution, there is also an important difference: the EDM assumes one distribution while SDT assumes two.

Conditioned Reinforcement

Conditioned reinforcement is the process whereby a previously neutral stimulus acquires the ability to strengthen responding associated with a primary reinforcer. The present research used a concurrent-chain procedure where subjects

chose between two alternatives that are maintained by the primary reinforcement (Grace & Savastano, 2000). The proponents of the concurrent chains procedure assumed that initial-link responses correspond to the relative conditioned reinforcement strength of the terminal-link stimuli (Autor, 1960; Herrnstein, 1964). Later researchers have also assumed that terminal-link stimuli act as conditioned reinforcers via Pavlovian conditioning (Grace, 1994). Existing steady-state models of choice, including DRT, HVA and CCM assume that terminal-link value is a prime determinant of initial-link response allocation (Grace, 1996). Response allocation by tracking the relative value of the terminal links is well-supported (Grace, 1994), and is referred to as the "value hypothesis" (Grace, 2002b). Thus, researchers have generally agreed that initial-link response allocation was determined by the value of the terminal-link stimuli, but disagreed about exactly how value should be calculated and how it influenced choice. The Contextual Choice Model (CCM: Grace, 1994) suggests conditioned reinforcement value is based on the relative terminal-link delay to reinforcement, while Delay Reduction Theory (DRT: Fantino, 1969) suggests conditioned reinforcement value is based on the relative reduction in delay to reinforcement, signalled by the terminal links, relative to the overall delay signalled by the initial links. The Hyperbolic Value Addition (HVA: Mazur, 2001) model assumes that conditioned reinforcement is based on the relative value added by entry into the terminal-link stimuli. DRT and HVA are similar in some respects because each is based on a difference comparison of the terminal-link stimuli with the prior initial link. Both use an overall amount of time then subtract an amount corresponding to the entry into the terminal link. The difference between them is that HVA calculates a hyperbolic value for delay to reinforcement for both overall and initial link delays, whereas DRT calculates the reduction from overall delay to reinforcement

using the actual terminal-link delays. CCM is different from both because it assumes terminal-link value is separate from the effects of temporal context, while HVA and DRT assert that value depends on temporal context. This is incorporated in CCM by having the expression for terminal-link value within parentheses, indicating construct boundaries (i.e., value) and independence, while exponents and remaining calculations indicate the effect of variables have secondary effect on the construction of the calculation. This method has been used to indicate the separate effect of temporal context on the calculation of terminal-link value (Grace, 1994). The independence of value and temporal context was supported by Grace and Savasatano (2000), who examined the value of terminal stimuli using occasional probe trials. Probe trials are measures of responding to stimuli previously presented in concurrent chains, without initial-link stimuli or reinforcement. Grace and Savasatano found probe choice was well described by the scheduled rates of reinforcement associated with each terminal link. The authors concluded that terminal-link value is independent of the temporal context of reinforcement, while context modulates the expression of learning in initial-link choice (Grace, 2002b). However, some researchers have come to different conclusions. Fantino and Romanovich (2007) reviewed several studies and suggested that elevated rates of responding in the terminal links of concurrent-chains schedules do not necessarily lead to greater preference in the initial link. They suggest that this is because the putative conditioned reinforcers in the terminal link are not correlated with a reduction in time to primary reinforcement or with an increase in value. Thus, there appears some disagreement about context and conditioned reinforcement value.

A study by Grace and Nevin (1999) casts further doubt on the validity of value as an explanatory construct for choice. They used a novel concurrent-chains

procedure which provided simultaneous measures of timing and choice. The sequence of the conditions they used was 1) a peak procedure, 2) peak procedure embedded in a concurrent-chains procedure, 3) a peak procedure, and 4) peak procedure embedded in a concurrent-chains procedure. Grace and Nevin found that timing on the peak procedure was the same as that with peak procedures embedded in the terminal-links in concurrent-chains. However, responding in the last condition appeared to reflect the terminal-link delays in the earlier concurrent-chains procedure rather than the terminal-link delays in the last concurrent-chains condition. Grace and Nevin suggested this is evidence contrary to the assumption that choice can be explained in terms of a fundamental timing process that determines both timing and choice (Gibbon et al., 1988).

By contrast, EDM accounts for choice in concurrent chains without relying on the construct of conditioned reinforcement. The EDM assumes a subject compares the value of the criterion with the just experienced terminal-link delay. The criterion is updated with the onset of stimuli correlated with reward. Thus, the EDM does not directly assume conditioned reinforcement value as the mechanism that determines choice. However, EDM does assume that terminal-link delay to reinforcement represents an important factor in determining initial-link response allocations. The distinction is that EDM compares the current terminal-link delay with a criterion based on combined histories, rather than a relative terminal-link value for each alternative. However, can the EDM account for the apparently paradoxical results of Grace and Nevin (1999)?

The EDM assumes that choice is determined by the discrimination of short or long delays based on the combined histories of initial- and terminal-link delays from both alternatives. However, in the third condition of Grace and Nevin (1999), the

initial-link stimuli were not presented. According to the EDM, the updating of initial-link response strengths cannot occur if the initial-link stimuli are not present. This would explain how preference was based on the contingencies in the earlier concurrent chains procedure and the slow change in behaviour in the last condition. Thus, according to EDM, the organism propensity to respond is developed by making a series of discriminations.

Summary of Experiments

The initial-link experiments confirmed the novel EDM prediction of a bitonic function as initial-link duration changed. This prediction occurs because relative response strength to the criterion is free to change with each alternative so predictions for each alternative can change at different rates. Predicted response allocation was found to peak at approximately the geometric mean of the experienced delays and mirrors temporal bisection experiments where the subjective midpoint between two values also approximates the geometric mean (Church & Deluty, 1977). The terminal-link experiments found the EDM to predict the 'terminal-link effect' (McEwen, 1972), and described greater VAC than LINOP (Grace, 2002a), which is also a linear-operator model of concurrent chains performance. Moreover, the EDM can describe both non-linear (categorical) and approximate linear (generalized matching) response allocations. In Experiment 4, for all pigeons, response allocation showed the descending series to be greater than the ascending series in the middle of the immediacy values. This indicates that response allocation appears to lag behind the changing immediacy, creating a separation between the ascending and descending response allocations.

The archival analyses showed that the EDM was able to describe initial-link response allocation across a wide range of published studies. In addition, residual analyses found that the EDM had no systematic deviations between obtained and predicted scores. Moreover, results suggested that terminal-link sensitivity was parameter invariant for EDM across different schedule types (VI or FI), but not for HVA and CCM. This gives the EDM a significant advantage over competing theories of steady-state choice. Coupled with the evidence of temporal context effects and the novel predictions of the bitonic effects, these results suggest that the EDM has promise as a comprehensive model of behavioural choice.

*Conclusion*s

The EDM belongs to a long line of successful linear operator models of choice acquisition. This method has been demonstrated to accurately describe the learning process in numerous experiments and has also been successfully applied to choice situations (Couvillion & Bitterman, 1985; Estes, 1951). However, a caveat regarding linear-operator models in general needs to be made. The mechanism that drives this model is very simple, and although it describes the general acquisition process across extended periods of time the assumption that a constant proportion determines the changes in predicted response strength is likely an oversimplification, and does not describe the variation in response allocation from moment to moment. Therefore, a linear operator model may need an additional process (or processes) that influences responding before we arrive at an accurate molecular model of choice. Nevertheless, the EDM has the advantage of parsimony over other models of choice that assume separate distribution histories for each alternative (Gallistel & Gibbon, 2000), and has

been designed to apply to concurrent chains, whereas other theories have been designed for simple concurrent schedules (Davison & Baum, 2000). Moreover, even models that allow for cognitive decision processes struggle with temporal context effects (Horner & Staddon, 1987), yet the EDM can predict these effects. Thus, the EDM appears to be a worthy contender amongst contemporary models for choice.

Researchers in the area of choice from different orientations have suggested that the value of competing stimuli is assessed independently of alternative stimuli while preference reflects an interaction between alternatives (Grace, 2002a; Stout & Miller, 2007). The common process these researchers assume that determines both response allocation and acquisition is Pavlovian conditioning. Grace (2002b) suggests that terminal-link conditioned reinforcement can be thought of in terms of a value hypothesis - where terminal-link stimuli acquire value in association with their proximity to primary reinforcement. Results from Grace and Nevin (1999) challenge this assumption and indicate that while timing remains congruent with delay to reinforcement, choice allocation appears to be controlled by prior conditioning. This result is difficult for theories that assume that both choice and timing processes are dependent on conditioned reinforcement. In contrast, the EDM assumes that the response strength of the choice alternatives changes as a function of the experienced delays, rather than relying on conditioned reinforcement. Thus, the probabilistic approach of the EDM, assuming a theoretical distribution of delay as part of the calculation of a criterion value and the probability of a short categorisation, provides an alternative explanation of choice phenomena that have traditionally been interpreted in terms of conditioned reinforcement.

Further Research

Further research could examine the generality of the EDM. For example, comparing EDM predictions with response allocations using different schedules such as: fixed or variable ratio schedules, probabilistic reinforcement, or comparing independent vs. dependent scheduling. Comparing the EDM with existing concurrent schedule models (Mazur, 1996) and (Davison & Baum, 2000), especially when applied to within session changes, would also test whether the model has generality to different experimental designs. Specifically, examining the nature of the EWMA as contingencies change within session would be informative as to whether the EDM or another process, such as reinforcer-accumulations (Davison and Baum, 2000), can describe response allocation in changing concurrent schedule components.

To test the generality of the EDM, two further assessments are presented. The first assessment asks, what is the prediction when the initial-link delays are progressively reduced, and the second, when the terminal-link delays are reduced? The following figure log response ratios predicted by the EDM are shown for this situation. The initial-link delay begins at 35s and is then progressively reduced to 0.1 whilst the terminal-link alternative is constant at 10s (left) and 12s (right). This assesses the robustness of the prediction of a bitonic effect using terminal link delays that are less discriminable. The EDM parameters Δ L, R and S are fixed at 0.2 while β is set at 0.3, $\sigma = 0.13$, $Min = 0.01$, $Max = 1.0$, and log b at 0.0. This results in the following prediction for response allocation:

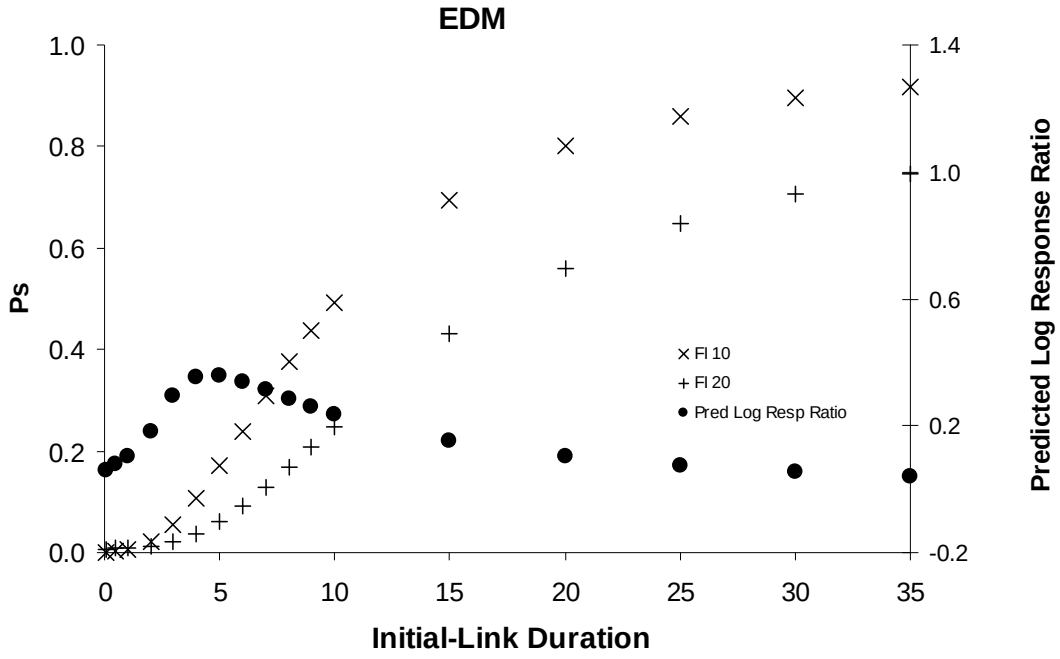


Figure 26. Log-initial-link response allocation (filled circles, right axis) predicted by the EDM as a function of initial-link duration. Data points indicated by x's and +'s show the probabilities of a terminal-link delay judged short ($P_{'s'}$, left axis) relative to the criterion for the FI 10 and FI 12s schedules, respectively.

In Figure 26 predicted log response ratios still show the bitonic pattern, although the peak of 0.4 log response ratio appears to have moved the left of the previous figure to occur at an initial-link duration of approximately 5 s, whilst the lowest log response ratio was approximately 0.1 log units obtained at 35 s. Therefore, reducing the difference between terminal-links does appear to affect the magnitude of the bitonic effect, but nonetheless, the effect is present.

Second, the EDM predicts that as schedules change a portion of predicted response allocation is carried over to the following schedule. This creates the possibility that response ratios could lag behind the changing immediacy values causing hysteresis effects. The following figure examines this situation for both the ascending and descending series – response rates change less to changes in smaller terminal-link delays in the ascending series and similarly for higher terminal-link delays in the descending series (Figure 27).

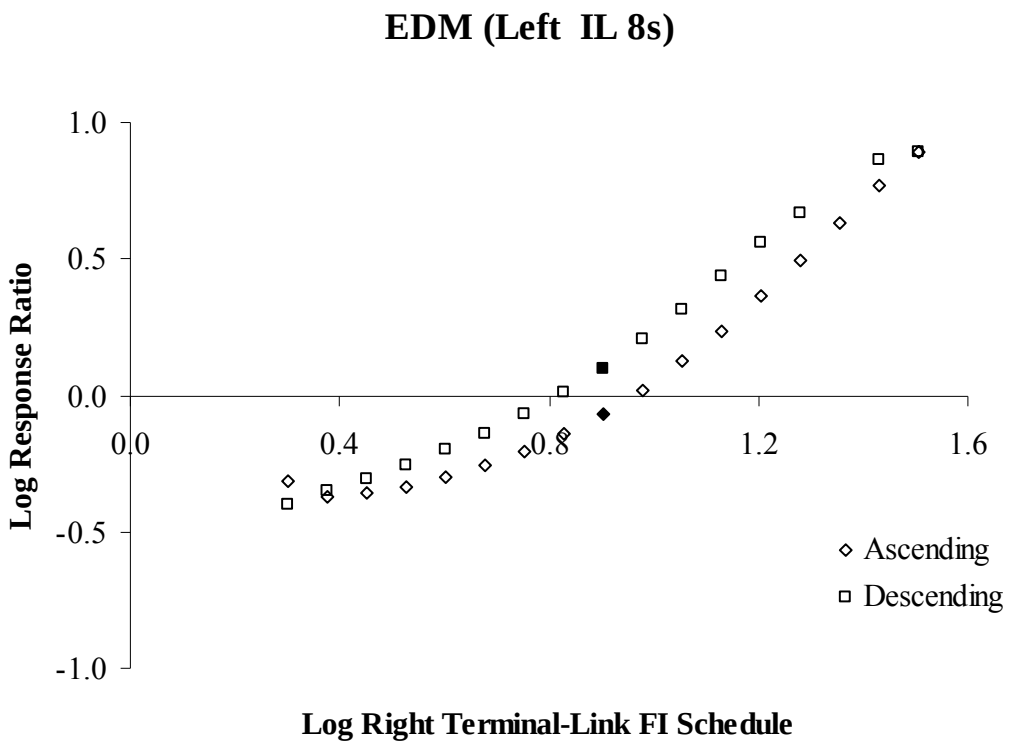


Figure 27. Predicted EDM Log initial response ratios as a function of terminal-link immediacy. Unfilled squares are descending delays whilst unfilled diamonds are ascending delays. The filled symbols are the mid-point 8 s for each series.

Where the initial-link is constant at 10s for each initial-link and the left terminal link is fixed at 8s, whilst the right terminal-link changes from 2 s to 32 s in either an ascending or descending sequence. EDM parameters Δ L, R and S are fixed at 0.2 while β is set at 0.3, $\sigma = 0.2$, $Min = 0.01$, $Max = 1.0$, and log b at 0.0. Thus, the EDM predicts hysteresis effects when one alternative changes. However, is this a persistent effect, does it exist at different delays? This question was examined when the difference between terminal-links was reduced and increased. The same initial-link delays and EDM parameters used in Figure 27 were used in the following simulation.

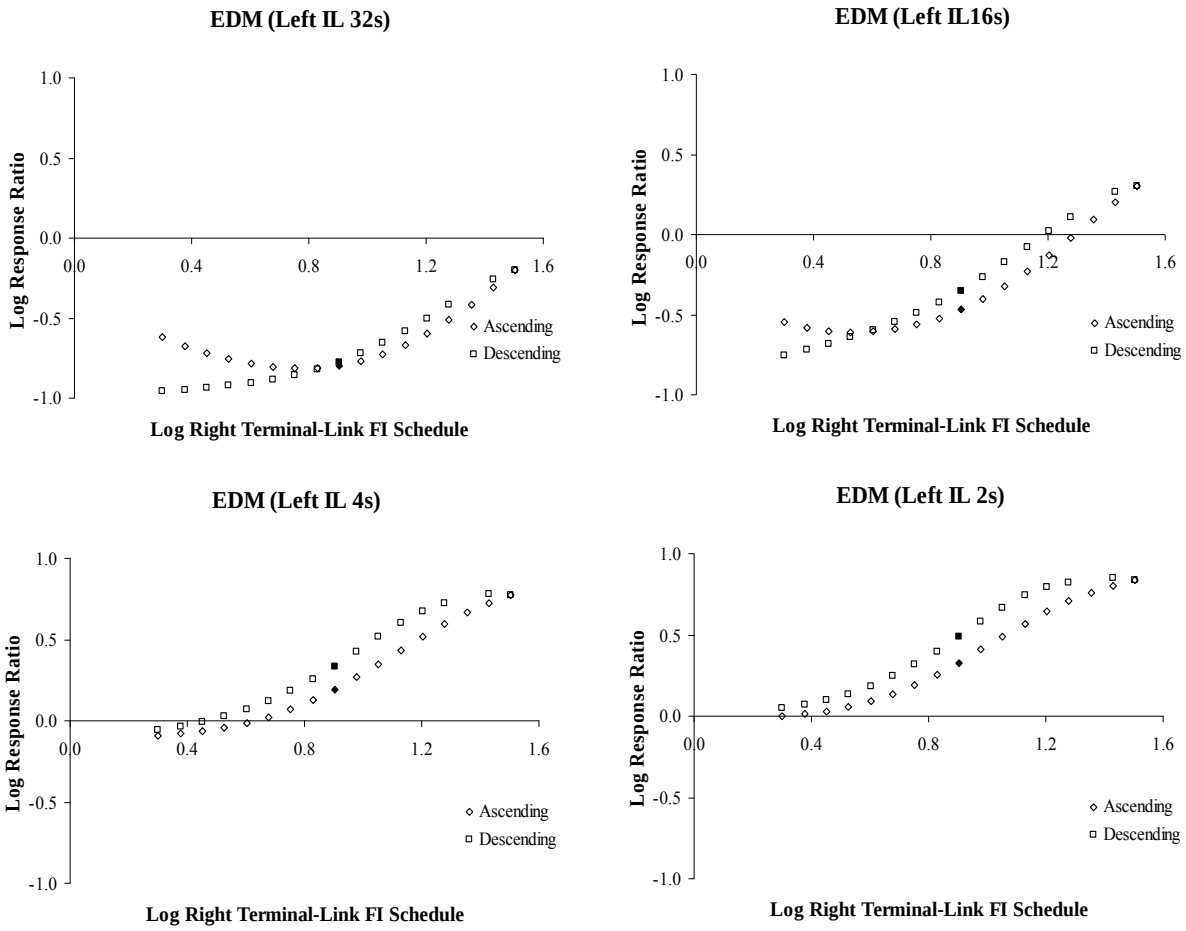


Figure 28. Predicted EDM Log initial response ratios as a function of terminal-link immediacy. Symbols are the same as previous figure.

Figure 28 shows hysteresis effects at different initial-link values: predicted log response ratios appear to persist at low and high terminal-link values. There is also an unusual prediction for the ascending series at longer terminal-link delays; values appear to become U shaped. This is the result of changing rates of predicted response allocation as the terminal-link delay changed. Increases in the right alternative response strength, and declines in left alternative predictions, as the delay increase approached 6.5 s, caused the predicted log response ratio to fall. However, delays greater than 6.5 s caused relatively greater declines in the right predicted response rates. This pattern occurred as a consequence of the corresponding right short categorisation probability falling, at a greater rate than the left. Therefore, predicted response ratio increased for terminal-link delays greater than 6.5 s. This unusual prediction typifies the effect of delays, probabilities of categorisations, and by extension, the criterion and sigma on EDM predicted response ratios. Further examination of the above predictions and the suggested changes to scheduling will test the generality of the EDM.

References

Autor, S. (1969). The strength of conditioned reinforcers as a function of frequency and probability of reinforcement. In D. Hendry (Ed.), *Conditioned Reinforcement* (pp. 127-419). Homewood, Illinois: The Dorsey Press.

Alsop, B., & Davison, M. (1988). Concurrent-chain performance: Effects of absolute and relative terminal-link entry frequency. *Journal of the Experimental Analysis of Behavior, 49*, 351-365.

Bailey, J., & Mazur, J. (1990). Choice behavior in transition: Development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior, 53,* 409-422.

Barofsky, I., & Hurwitz, D. (1968). Within ratio responding during fixed ratio performance. *Psychonomic Science, 11,* 263-64.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22,* 231-242.

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 32,* 269-281.

Baum, W. M., & Rachlin, H. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12,* 861-874.

Brown, P. L., & Jenkins, H. M. (1968). Autoshaping of the pigeon's key peck. *Journal of the Experimental Analysis of Behavior, 11*, 1-8.

Bush, R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review, 58,* 313-323.

Bush, R., & Mosteller, F. (1955). *Stochastic Models for Learning*. New York: Wiley.

Chung, S., & Herrnstein, R. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 10,* 67-74.

Church, R., & Deluty, M. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 216-228.

Couvillon, P., & Bitterman, M. (1985). Analysis of choice in honeybees. *Animal Learning & Behavior, 13*, 246-252.

Davison, M. C. (1976). Preference for fixed-interval schedules: Effects of unequal initial links, *Journal of the Experimental Analysis of Behavior, 25*, 371-376.

Davison, M. C. (1983). Bias and sensitivity to reinforcement in a concurrent schedule. *Journal of the Experimental Analysis of Behavior, 40,* 15-34.

Davison, M. C. (1987). The analysis of concurrent-chain performance. In M. Commons, J. Mazur, J. Nevin, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior, 5: The effect of delay and of intervening events on reinforcement value* (pp. 225-241). Hillsdale, NJ: Erlbaum.

Davison, M. C., & Baum, W. (2000). Choice in a variable environment: Every reinforcer counts. *Journal of the Experimental Analysis of Behavior, 74*, 1-24.

Davison, M. C., & Hunter, I. (1979). Concurrent schedules: Under-matching and control by previous experimental conditions. *Journal of the Experimental Analysis of Behavior, 32,* 233-244.

Davison, M. C., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behaviour, 13*, 77-84.

Davison, M. C., & Jones, B. M. (1998). Performance on concurrent variable-interval extinction schedules. *Journal of the Experimental Analysis of Behavior, 69*, 49-57.

Davison, M. C., & McCarthy, D. (1988). *The Matching Law: A research review*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davison, M. C., & Temple, A. (1973). Preference for fixed-interval schedules: An alternative model. *Journal of the Experimental Analysis of Behavior, 20*, 393-403.

Devenport, L. D., & Devenport, J.A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behavior, 47*, 787-802.

Dreyfus, L. (1991). Local shifts in relative reinforcement rate and time allocation on concurrent schedules. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 486-502.

Duncan, B., & Fantino, E. (1970). Choice for periodic schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 14*, 73-86.

Dunn, R., & Fantino (1982). Choice and the relative immediacy of reinforcement. *Journal of the Experimental Analysis of Behavior, 38*, 321-326.

Estes, W. (1950). Toward a statistical theory of learning. *Psychological Review, 57*, 94-107.

Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior, 12*, 723-730.

Fantino, E. (2001). Context: A central context. *Behavioural Processes, 54*, 95-110.

Fantino, E., & Davison, M. C. (1983). Choice: Some quantitative relations. *Journal of the Experimental Analysis of Behavior, 40*, 1-13.

Fantino, E., Preston, R., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior, 60,* 159-169.

Fantino, E., & Romanowich, P. (2007). The effect of conditioned reinforcement rate on choice: A review. *Journal of the Experimental Analysis of Behavior, 87,* 409-421.

Fantino, E., & Royalty, P. (1987). A molecular analysis of choice on concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior, 33*, 27-37.

Fantino, E., Squires, N., Delbruck, N., & Peterson, C. (1972). Choice behavior and the accessibility of the reinforcer. *Journal of the Experimental Analysis of Behavior, 18*, 35-43.

Field, D. P., Tonneau, F., Ahearn, W., & Hineline, P. N. (1996). Preference between variable-ratio and fixed ratio schedules: Local and extended relations. *Journal of the Experimental Analysis of Behavior, 66*, 283-295.

Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5*, 529-530.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107,* 289-344.

Gallistel, C. R., & Gibbon, J. (2002). *The Symbolic Foundations of Conditioned Behavior.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of the Experimental Analysis of Behavior, 27*, 354-372.

Gentry, D. G., & Marr, J. M. (1980). Choice and reinforcement delay, *Journal of the Experimental Analysis of Behavior, 33*, 27-37.

Gibbon, J. (1971). Scalar timing and semi-markov chains in free-operant avoidance. *Journal of Mathematical Psychology, 8*, 109-138.

Gibbon, J. (1977). Scalar expectancy theory and weber's law in animal timing. *Psychological Review, 84*, 279-325.

Gibbon, J., & Balsalm, P. (1981). Spreading associations in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and Conditioning Theory* (pp. 219-253). New York: Academic Press.

Gibbon, J., Baldock, M., Lucurto, C., Gold, L., & Terrace, H. (1977). Trial and intertrial duration in autoshaping. *Journal of Experimental Psychology, 3*, 264-284.

Gibbon, J., Church, R., Fairhursh, S., Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review, 95*, 102-144.

Grace, R. C. (1994). A contextual model of concurrent chains choice. *Journal of the Experimental Analysis of Behavior, 61*, 113-129.

Grace, R. C. (2002a). Acquisition of preference in concurrent chains: Comparing linear-operator and memory-representational models. *Journal of Experimental Psychology: Animal Behavior Processes, Vol 28,* 257-276.

Grace, R. C. (2002b). The value hypothesis and acquisition of preference in concurrent chains. *Animal Learning & Behavior, 30*, 21-33.

Grace, R. C. (2004). Temporal context in concurrent chains I: Terminal-link duration. *Journal of the Experimental Analysis of Behavior, 81,* 215-237.

Grace, R. C., Berg, M. E., & Kyonka, E. (2006). Choice and timing in concurrent chains: Effects of initial-link duration. *Behavioural Processes, 71,* 188-200.

Grace, R. C., & Bragason, O. (2004). Does the terminal-link effect depend on duration or reinforcement rate? *Behavioural Processes, 67*, 67-79.

Grace, R. C., Bragason, O., & McLean, A. (2003). Rapid acquisition of preference in concurrent chains. *Journal of the Experimental Analysis of Behavior, 80,* 235-252.

Grace, R. C., & McLean, A. (2006). Rapid acquisition in concurrent chains: Evidence for a decision model. *Journal of the Experimental Analysis of Behavior, 85,* 181-202.

Grace, R. C., & Nevin, J. A. (1999). Timing and choice in concurrent chains. *Behavioural Processes, 45*, 115-127.

Grace, R. C., & Savsatano, H. I. (2000). Temporal context and conditioned reinforcement value. *Journal of Experimental Psychology: General, 129*, 427-443.

Green, L. Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin, 130*, 769-792.

Green, D. M., Swetz, J. A. (1966). *Signal detection theory and psychophysics.* Wiley, New York.

Greenwood, M. R. C., Quartermain, D., Johnson, P. R., Cruce, J. A. F. & Hirsch, J. (1974). Food motivated behavior in genetically obese and hypothalamic-hyperphagic rats and mice. *Physiology & Behavior, 13*, 687-92.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.

Herrnstein, R. J. (1964). Aperiodicity as a factor in choice. *Journal of the Experimental Analysis of Behavior, 7,* 179-182.

Herrnstein, R. J., & Heyman, G. M. (1979). Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior, 31,* 209-223.

Herrnstein, R. J., & Vaughan, W. (1980). Melioration and behavioural allocation. In J. E. R. Staddon (Ed.), *Limits to Action: The allocation of individual behaviour (pp143-176)*. New York: Academic Press.

Horner, J. M., & Staddon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Process, 15*, 59-92.

Hunter, I., & Davison, M. C. (1985). Determination of a behavioral transfer function: White-noise analysis of session-to-session response-ratio dynamics on current VI VI schedules. *Journal of the Experimental Analysis of Behavior, 43*, 43-59

Jenkins, H. M., & Moore, B. R. (1973). The form of the autoshaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior, 20*, 163-181.

Kelsey, J. E., & Allison, J. (1976). Fixed-ratio lever pressing by VMH rats: Work vs accessibility of sucrose reward. *Physiology & Behavior, 17*, 749-54.

Killeen, P. R. (1970). Preference for fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 14*, 127-131.

Killeen, P. R. (1972). The matching law. *Journal of the Experimental Analysis of Behavior, 17,* 489-495.

Killeen, P. R. (1981). Incentive theory. *Nebraska Symposium on Motivation, 29*, Lincoln: University of Nebraska Press.

Killeen, P. R. (1982). Incentive theory II: Models for choice. *Journal of the Experimental Analysis of Behavior, 38*, 217-232.

Killeen, P. R. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences, 17,* 105-172.

Kyonka, E. G. E., & Grace, R. C. (2007). Rapid acquisition of choice and timing in pigeons. *Journal of Experimental Psychology. Animal Behavior Processes, 33*, 392-408

Kyonka, E. G. E., & Grace, R. C. (2008). Rapid acquisition of preference in concurrent chains when alternatives differ on multiple dimensions of reinforcement. *Journal of the Experimental Analysis of Behavior*, *89*, 49-69.

LaBounty, C, & Reynolds, G. (1973). An analysis of response and time matching to reinforcement in concurrent ratio-interval schedules. *Journal of the Experimental Analysis of Behavior, 19*, 155-166.

Lea, S. E. G., & Dow, S. (1984). The integration of reinforcements over time. *Annals of the New York Academy of Sciences*, *423*, 269-277.

MacEwen, D. (1972). The effects of terminal-link fixed-interval and variable-interval schedules on responding under concurrent-chained schedules. *Journal of the Experimental Analysis of Behavior, 18*, 253-261.

Mazur, J. E. (1982). A molecular approach to ratio schedule performance. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior: Matching and maximising accounts* (pp79-110). Cambridge, USA: Ballinger.

Mazur, J. E. (1984). Tests for an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal behavior processes, 10*, 426-351.

Mazur, J. E. (1992). Choice behavior in transition: Development of preference with ratio and interval schedules. *Journal of Experimental Psychology: Animal behavior processes, 18,* 364-378.

Mazur, J. E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning & Behavior, 23,* 93-103

Mazur, J. E. (1996). Past experience, recency, and spontaneous recovery in choice behavior, *Animal Learning and Behavior, 24,* 1-10.

Mazur, J. E. (1997). Effects of rate of reinforcement and rate of change on choice behaviour in transition. *The Quarterly Journal of Experimental Psychology, 50B*, 111-128.

Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review, 108*, 96-112.

Mazur, J E. (2004). Varying initial-link and terminal-link durations in concurrent-chains schedules: a comparison of three models. *Behavioural Processes, 66*, 189-200.

Mazur, J. E. (2005). Exploring a concurrent-chains paradox: Decreasing preference as an initial-link is shortened. *Journal of Experimental Psychology: Animal behavior processes, 31*, 3-17.

Mazur, J. E., Blake, N., & McManus, C. (2001). Transitional choice behavior in concurrent-chains schedules. *Behavioral Processes, 53,* 171-180.

Mazur, J. E., & Ratti, T. (1991). Choice behavior in transition: Development of preference in a free-operant procedure. *Animal Learning & Behavior, 19,* 241-248.

Myerson, J., & Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior, 49,* 291-302.

Myerson, J., & Miezin, F. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review, 87,* 160-174.

Navarick, D. J., & Fantino, E. (1976). Self-control and general models of choice, *Journal of Experimental Psychology: Animal Behavior Processes, 2*, 75-87.

Nevin, J. A. (1973). Conditioned reinforcement. In J. A. Nevin & G. S. Reynolds (Eds.), *The Study of Behavior: Learning, motivation, emotion, and instinct* (pp. 154-198). Glenview, IL: Scott, Foresman.

Nevin, J. A. (1984). Quantitative analysis. *Journal of the Experimental Analysis of Behavior, 42*, 421-434.

Nevin, J. A. (1992). Behavioral contrast and behavioral momentum. *Journal of Experimental Psychology: Animal behavior processes, 18*, 126-133.

O'Daly, M., Meyer, S., & Fantino, E. (2005). Value of conditioned reinforcers as a function of temporal context. *Learning and Motivation, 36*, 42-59.

Omino, T., & Ito, M. (1993). Choice and delay of reinforcement: Effects of terminal-link stimulus and response conditions. *Journal of the Experimental Analysis of Behavior, 59*, 361-371.

Pavlov, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109*, 472-491.

Powell, R. W. (1968). The effect of small sequential changes in fixed-ratio size upon the post-reinforcement pause. *Journal of the Experimental Analysis of Behavior, 11*, 589-93.

Preston, R. & Fantino, E. (1991). Conditioned reinforcement value and choice. *Journal of the Experimental Analysis of Behavior, 55,* 155-175.

Rachlin, H. (1971). On the tautology of the matching law. *Journal of the Experimental Analysis of Behavior, 15*, 249-251.

Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. (1976). Economic demand theory and psychological studies of choice. In G. Bower (Ed.), *The Psychology of Learning and Motivation. (pp 129-154).* New York: Academic Press.

Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical Conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Process, 7*, 242-268.

Schofield, G., & Davison, M. (1997). Nonstable concurrent choice in pigeons. *Journal of the Experimental Analysis of Behavior, 68*, 219-232.

Silverberg, A., Ziriax, J. M. (1982). The inter-changeover time as a molecular dependent variable in concurrent schedules. In M. L. Commons, R. J., Herrnstein, & H. Rachlin (Eds.), *Quantative Analysis of Behaviour, Vol 2: Matching and maximising accounts* (pp. 131-151). Cambridge, MA: Ballinger.

Skinner, B. F. (1986). Some thoughts about the future. *Journal of the Experimental Analysis of Behavior, 45,* 229-235.

Spetch, M. L., & Dunn, R. (1987). Choice between reliable and unreliable outcomes: Mixed percentages – reinforcement in concurrent chains. *Journal of the Experimental Analysis of Behavior, 47, 57-72.*

Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior, 15*, 27-38.

Staddon, J. E. R. (1977). On Herrnstein's equation and related forms. *Journal of the Experimental Analysis of Behavior, 28,* 163-170.

Staddon, J. E. R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior, 49*, 303-320.

Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review, 114*, 759-783.

Stubbs, D. A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental of Behavior, 11,* 223-238.

Sutton, N. P., Grace, R. C., McLean, A. P., & Baum, W. M. (2008). Comparing the generalized matching law and contingency discriminability model as accounts of concurrent schedule performance using residual meta-analysis. *Behavioural Processes, 78,* 224-230.

Terrace, H. S., Gibbon, J., Farrell, L., & Baldock, M. D. (1975). Temporal factors influencing the acquisition and maintenance of an autoshaped keypeck. *Animal Learning and Behavior, 3*, 53-62.

Thorndike, E. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement, 2*, 1-09.

Wardlaw, G. R. & Davison, M. C. (1974). Preference for fixed-interval schedules: Effects of initial-link length. *Journal of the Experimental Analysis of Behavior, 21*, 331-340.

Williams, B. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review, 1*, 457-475.

Williams, B., & Dunn, R. (1991). Substitutability between conditioned and primary reinforcers in discrimination acquisition. *Journal of the Experimental Analysis of Behavior, 55,* 21-35.

Williams, B., & Fantino, E. (1978). Effects on choice of reinforcement delay and conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 29*, 77-86.